

THE ROLE AND
METHODOLOGY

OF Classification
in Psychiatry and
Psychopathology

LIBRARY

APR 11 1968

NATIONAL INSTITUTES OF HEALTH

Conference on
**THE ROLE AND
METHODOLOGY**

**OF Classification
in Psychiatry and
Psychopathology**

Washington, 1965,

Proceedings of a Conference held in Washington, D.C., November 1965
under the auspices of

The American Psychiatric Association and
The Psychopharmacology Research Branch
National Institute of Mental Health

Edited by
Martin M. Katz, Ph. D.
Jonathan O. Cole, M.D.
Walter E. Barton, M.D.

U.S. DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE
Public Health Service
National Institute of Mental Health, Chevy Chase, Md. 20203

Acknowledgments

The editors owe special thanks to Samuel Lyerly, Bureau of Social Science Research, and Samuel Greenhouse, National Institute of Mental Health, for contributing freely of their time in the initial development of the aims, and in advising on the selection of critical topics within mathematical statistics and psychometrics to be covered at the conference.

The editors also express their appreciation to Robert L. Robinson, Public Information Officer, American Psychiatric Association for his invaluable guidance throughout in planning and managing the conference.

PUBLIC HEALTH SERVICE PUBLICATION No. 1584

For sale by the Superintendent of Documents, U.S. Government Printing Office
Washington, D.C. 20402 - Price \$4.50

Preface: THE RATIONALE AND GENERAL STRUCTURE OF THE CONFERENCE

Martin M. Katz, Ph. D.

Jonathan O. Cole, M.D.

Walter E. Barton, M.D.

The Conference on the Role and Methodology of Classification in Psychiatry and Psychopathology was the result of a joint effort by the Psychopharmacology Research Branch of the National Institute of Mental Health and the American Psychiatric Association, to organize a comprehensive and, hopefully, definitive conference on the basic problem of classification in psychopathology and psychiatry.

It grew out of a realization that a continually increasing body of valuable research is accumulating which calls for reexamination of the bases of psychiatric diagnosis, and that many new approaches to typing or classifying patients are being investigated for their clinical or research utility and their promise for broadening our understanding of psychopathology. It was hoped that the conference would insure both that future research on classification would take into account relevant advances in mathematical statistics and that distinguished statisticians and behavioral scientists would get a full picture of the specific problems facing psychiatry.

The major aims of the conference can best be described by briefly reviewing its rationale, and the program structure which evolved from discussions of the editors with a specially selected NIMH-APA Conference Planning Committee.¹

The editors and the planning committee were aware that over the last 10 years, an increasing number of investigators have been involved in attempts to develop new systems for classifying mental disorders. They attributed the increase in research activity in this area to several sources.

First, there has been a growing awareness among clinicians and researchers that the lack of an objective and reliable system for diagnosing the mental disorders limits the carrying out of important basic research in the field of psychopathology. Secondly, the field has witnessed a sharp rise over the last few years in the number and diversity of treatments for the mentally disturbed, and thus, the interest in the problem of treatment specificity; i.e., the problem of which type of treatment is more effective for which type of patient, has intensified and become more compelling. These developments have occurred in the context of an increasing dissatisfaction with the traditional system of psychiatric diagnosis; a dissatisfaction which is based on both the demonstrated technical deficiencies of the system and on its failure to be very useful in prognosis.

The planning committee, aware of the sheer size and complexity of the classification problem, did not view the conference as an attempt to solve the psychiatric diagnosis problem. What had been observed was that the need for new objective systems had become more pressing and that the increase in research activities in this area had forced a number of unresolved issues into somewhat sharper focus.

¹ See app. 1.

There is, for example, a good deal of confusion within and between such related fields as clinical psychiatry, epidemiology, and psychopathology, both as to the meaning and function of diagnosis, and as to the place of diagnosis in various types of research activity in these fields. In addition, recent attempts to develop new systems have resulted in new problems of a highly technical, statistical nature with regard to how a new classification system might be developed.

The committee thought that by bringing together people who were concerned with the varying roles of diagnosis, investigators who were active in research in this field, and experts on some of the technical, statistical problems confronting the field, it might be possible to clear away some of this confusion and make some progress on certain core issues of research in the area.

The issues deemed most important by the committee are reflected in the organization of the program; they are stated briefly here as a way of clarifying the aims of the conference.

It is apparent that classification in the field of psychopathology may mean different things to researchers and to clinicians, and its meaning appears to be tied to the role or function it is expected to play in the investigator's research or his practice.

Classification can, therefore, have several functions and these functions are not necessarily compatible. The differences in purpose are apparent when we contrast the role of diagnosis in clinical psychiatry where the concern is with how this system will help in planning treatment, with its role in psychopathology, where the interest may be in how classification will facilitate the identification and description of certain etiological factors.

The first set of papers attempts to place this problem in perspective. The result when one compares the formulations of the epidemiologist, the hospital psychiatrist, and the psychopathologist is to force the question of whether a single comprehensive

system is really feasible in this field. This issue of whether one or several systems are required is one which the conferees were asked to consider.

The varied functions of diagnosis also make us aware that there is less agreement about what constitute the fundamental bases for classification—that is, which aspects of human functioning, which human characteristics, are fundamental to the distinctions between types of patients. The conferees concluded that it was important not to limit inquiry to the fields of psychopathology and personality.

For this reason, experts from other disciplines, for example, sociology, biology, and human development were asked to look at these problems from their own vantage points, and to review developments in their own fields which may be useful in determining the nature of the information which is necessary to a classification system.

There follows, then, the question as to the proper psychometric and mathematical-statistical techniques for measuring and combining these characteristics, this information, and discriminating and separating out types of people. A section of the program was devoted to the review of recent technical advances in these fields.

In the third section, a number of attempts to develop new systems, new ways of organizing what is known about the mental disorders were presented. This group of investigators have in common the fact that they have all had practical experience with the array of problems which underlie the development of new systems. As such, they have been forced to make decisions at each step of the process with regard to problems which were as yet unresolved. These investigators had temporarily resolved the problems in the area in different ways. For example, they differed in—

1. theoretical orientation, that is, on what they considered to constitute the relevant phenomena to be measured;

2. psychological and clinical methods chosen for measuring these phenomena,

and in the psychometric approach selected to refine and to establish the validity of these methods;

3. kinds of human populations they selected to study;

4. statistical methods selected for determining whether qualitatively different types exist in a population; and

5. approaches to validating a new typology.

In a sense, the above are the core issues with which the planners were concerned and the ones on which it was hoped that this conference would be able to shed some light.

In limiting the scope of the conference, the committee hoped to make more progress in deepening our understanding of how new systems are being or might be developed than in resolving the more controversial theoretical issues underlying the content of such systems.

In this respect, the conference did appear to effect a partial integration of theory, practice, and method in the field of classification. It is hoped that the conference proceedings and recommendations can serve to provide a firmer base for future research of any type in this very complex but highly important problem area.

Contents

	Page		Page
Preface: The Rationale and General Structure of the Conference.....	iii	B. The Role of Classification in Research in Psychopathology	
<i>Martin M. Katz, Ph. D.</i>		Epidemiology and Medical Care	
<i>Jonathan O. Cole, M.D.</i>		Statistics	76
<i>Walter E. Barton, M.D.</i>		<i>Ernest M. Gruenberg, M.D.,</i>	
Introduction and Overview of the Conference	1	<i>Dr. P.H.</i>	
<i>Jonathan O. Cole, M.D. and</i>		Classification of Mental Disorders for Epidemiologic and Medical Care Purposes: Current Status, Problems, and Needs	99
<i>Martin M. Katz, Ph. D.</i>		<i>Morton Kramer, Sc. D.</i>	
The History of Classification in the Medical Sciences	11	The Role of Classification in the Development of the Science of Psychopathology with Particular Reference to Research	116
<i>Owsei Temkin, M.D.</i>		<i>David Shakow, Ph. D.</i>	
Discussion	20	Classification in Research on the Prediction of Response to Specific Treatments in Psychiatry	143
THE VARYING ROLES OF CLASSIFICATION IN PSYCHIATRY AND PSYCHOPATHOLOGY		<i>Jonathan O. Cole, M.D.</i>	
A. The Role of Classification in Psychiatric Practice		Discussion: The Role of Classification in Research in Psychopathology	147
The Role of Classification in Hospital Psychiatry	29	C. The Role and Problems of Classification in Related Disciplines and Their Relevance to Problems in Psychiatry and Psychopathology	
<i>Henry Brill, M.D.</i>		The Role of Classification in Personality Theory	155
The Role of the Classification System in Outpatient Psychiatry	35	<i>George A. Kelly, Ph. D.</i>	
<i>Elmer A. Gardner, M.D.</i>		Abnormal Psychology: Nosology Adrift	164
A Plan for Refining the Nosology of Mental Illness	53	<i>Kenneth R. Hammond, Ph. D.</i>	
<i>Paul T. Wilson, M.D.</i>			
The Role of Classification in Psychoanalytic Practice	62		
<i>Harry I. Weinstock, M.D.</i>			
Discussion: The Role of Classification in Psychiatric Practice	72		

	Page		Page
Some Implications of Classification in Sociology for Problems of Classification in Psychiatry and Psychopathology	168	Studies of Quantitative Approaches to Psychiatric Classification	277
<i>John A. Clausen, Ph. D.</i>		<i>John E. Overall, Ph. D. and Leo E. Hollister, M.D.</i>	
Discussion of Papers by Brill, Weinstock, Gleser, Overall and Hollister, and Stein and Neulinger in the Light of Research in Human Development	173	A Phenomenological Typology of Schizophrenia	300
<i>Lois B. Murphy, Ph. D.</i>		<i>Martin M. Katz, Ph. D.</i>	
Problems in Psychiatric Nosology From the Viewpoint of the Biological Sciences	190	Discussion: Descriptive and Phenomenological Approaches	
<i>Seymour S. Kety, M.D.</i>		Formal Discussion of Papers:	
		<i>Ardie Lubin, Ph. D.</i>	322
		<i>Samuel W. Greenhouse, Ph. D.</i>	328
		<i>Heinz E. Lehmann, M.D.</i>	330
		Open Discussion	344
ISSUES IN THE METHODOLOGY AND STATISTICS OF CLASSIFICATION		B. Systems Based on Patterns of Psychological Test Performance	
On the Meaning of Discrimination, Classification, Mixture, and Clustering in Statistics	197	Biometric Assessment of Mental Patients	353
<i>Samuel W. Greenhouse, Ph. D.</i>		<i>Joseph Zubin, Ph. D.</i>	
Quantifying Similarity Between People	201	Patterns of Intellectual Functioning and Their Implications for the Dynamics of Behavior	377
<i>Goldine C. Gleser, Ph. D.</i>		<i>David R. Saunders, Ph. D. and John W. Gittinger, M.S.</i>	
Multidimensional Representation of Similarity Structures	212	A Typology of Self-Descriptions.....	390
<i>Warren S. Torgerson, Ph. D.</i>		<i>Morris I. Stein, Ph. D. and John Neulinger, Ph. D.</i>	
A Survey of Some Empirical Clustering Procedures	220	Discussion: Systems Based on Patterns of Psychological Test Performance	
<i>Samuel B. Lyerly, Ph. D.</i>		Formal Discussion of Papers:	
Discrimination Among Groups and Assigning New Individuals	229	<i>Robert E. Harris, Ph. D.</i>	404
<i>C. Radhakrishna Rao, Sc. D., F.R.S.</i>		<i>Edward W. Forgy, Ph. D.</i>	410
Discussion: Issues in the Methodology and Statistics of Classification	240	Open Discussion	415
		C. Systems Based on Premorbid History, Course of Illness, or Reaction to Treatment	
THE CRITICAL REVIEW AND DISCUSSION OF THE METHODOLOGY INVOLVED IN A SERIES OF ATTEMPTS TO DEVELOP NEW TYPOLOGIES		Process and Reactive Schizophrenia:	
A. Descriptive and Phenomenological Approaches		Some Conceptions and Issues	419
The Phenomena of Depressions	249	<i>Norman Garnezy, Ph. D.</i>	
<i>Roy R. Grinker, Sr., M.D. and Jum C. Nunnally, Ph. D.</i>		Diagnosis and Pattern of Reaction to Drug Treatment: Clinically Derived Formulations	466
A Typology for Functional Psychotics..	261	<i>Donald F. Klein, M.D.</i>	
<i>Maurice Lorr, Ph. D.</i>		Patterns of Reaction to Drug Treatments Derived Through Multivariate Procedures	483
		<i>Dean J. Clyde, Ph. D.</i>	

	Page
Discussion: Systems Based on Premor- bid History, Course of Illness, or Reaction to Treatment	491
Formal Discussion of Papers:	
<i>Heinz E. Lehmann, M.D.</i>	491
<i>C. James Klett, Ph. D.</i>	493
Open Discussion	500
D. Systems Which Attempt To Encompass Variables From Several of the Significant Dimensions of Functioning in Mental Illness	
Typology of Schizophrenia Based on Multidisciplinary Observational Vectors	507
<i>Nils B. Mattsson, LL.M. and Ralph W. Gerard, M.D.</i>	
Neurophysiological Response Strategies in the Classification of Mental Illness	535
<i>Max Fink, M.D.</i>	
Discussion: Systems Which Attempt To Encompass Variables From Several of the Significant Dimensions of Functioning in Mental Illness	540
Formal Discussion of Papers:	
<i>Eugene Laska, Ph. D.</i>	540
<i>Max Hamilton, M.D.</i>	542
Open Discussion	543

PERSPECTIVES ON THE CONFERENCE

<i>Max Hamilton, M.D.</i>	551
<i>Joseph Zubin, Ph. D.</i>	556
<i>C. Rahdakrishna Rao, Sc. D., F.R.S.</i>	559

REFLECTIONS ON THE MAJOR CONFERENCE ISSUES

563

<i>Martin M. Katz, Ph. D and Jonathan O. Cole, M.D.</i>	
---	--

Appendix I.—American Psychiatric Association-National Institute of Mental Health Conference Planning Committee	569
---	-----

Appendix II.—Program. Conference on the Role and Methodology of Classification in Psychiatry and Psychopathology	570
---	-----

Appendix III.—Alphabetical Roster. Conference on the Role and Metho- dology of Classification in Psychiatry and Psychopathology	573
--	-----

Index	577
-------------	-----

Introduction and Overview of the Conference

*Jonathan O. Cole, M.D., and
Martin M. Katz, Ph. D.*

The general organization of this volume has been summarized in the preface. It should be noted that the major papers to be found in this volume were prepared well in advance of the conference and were circulated to the participants. The conference itself, therefore, consisted of active discussion of the individual papers plus a free-ranging discussion of a number of issues concerning classification and research on classification.

The volume offers a number of things to any reader. To the reader relatively sophisticated in psychiatric research and its statistical methodology it offers an up-to-date overview of a number of major research programs which have applied statistical or other approaches to the classification of psychiatric patients into subtypes. In some instances this is the first major report of the results of a research program; for other programs the chapters in this volume will provide a concise overview of work accomplished to date, how the work was begun, its logic, its methods of procedure, and its results to date. To the less statistically sophisticated reader, the book provides not only an excellent chance to familiarize himself with this work but also a wide variety of excellent less technical chapters in areas of major interest to the practicing psychiatrist, the psychiatric resident, the medical student, or the clinical psychologist.

Some chapters which, in the opinion of the editors, seemed likely to be particularly

rewarding to the clinical reader are Dr. Gruenberg's excellent presentation of the role of epidemiological research in psychiatry, Dr. Garnezy's superb review of the process-reactive dimension in schizophrenia, and Dr. Shakow's fascinating description of the major research effort directed to unraveling the problems of schizophrenia carried out at the Worcester State Hospital 20 years ago, its methodological advances and the implications of this work for the future. Dr. Weinstock's excellent paper, presenting results of a survey of psychoanalytic practice completed a number of years ago, is published here for the first time, as a chapter in this volume. A particularly lucid series of papers on the role and purpose of statistics in research in this area are included. These are written in such a way as to make their content meaningful to the nonstatistician, who hopefully will emerge from the reading of the chapters by Drs. Greenhouse, Gleser, Rao, Torgerson, and Lyerly with a clear and useful notion of the purposes to which individual statistical approaches are suited, their limitations as well as their assets.

In the discussion sections (which contain only a selected portion of the total transcribed discussions), one will get the flavor of some vigorous arguments between clinicians and nonclinicians, and even between clinicians on the role of classification and the appropriateness of various methods to the development of better classification systems. The discussion sections are strong

on historical perspective. To those who feel overwhelmed by rating scales and statistics, Dr. Lehmann provides a spirited offense from the clinician's standpoint attacking possible defects in current research utilizing these approaches.

Extremes represented at this conference ranged widely. Dr. Lawrence Kubie's general position is that the complexity of events leading to the development of a psychiatric illness, and the complexity of the factors modulating and modifying its ever-changing course over time as well as the multiple factors, operating on both patient and therapist-observer during any interview situation, are overly rich in important and poorly understood details. Therefore, any attempt at classification based on our current understandings of patients and their psychopathology must of necessity be inadequate and premature. The extreme opposite position was represented by Dr. Ball, a computer expert from California. He expressed the lively faith that the stickier and more complicated difficulties in describing patients, understanding the interrelationships between various aspects of psychopathology or psychiatric history and diagnosis or subtype, could be readily resolved if psychiatrists and statisticians could be wedded through the computer, assuming the computer could be utilized as a flexible exploratory research instrument. These represent the extremes of optimism and pessimism regarding the utility of statistics and computers in clarifying problems of psychiatric classification.

The conference was also productive of wonderful historical quotes from a number of very distinguished people. One example from D. Hack Tuke seems appropriate here: "The wit of man has rarely been more exercised than in the attempt to classify the morbid mental phenomena covered by the term 'insanity.' The result has been disappointing."

To demonstrate that statements made in the present conference are also quotable

we would note a comment made by Dr. Henry Brill, "It is very hard to classify one's ignorance in a satisfactory fashion."

As the reader will note, it is a good deal easier to be negativistic about classification than to be optimistic. At least, derogatory criticisms seem a good deal more quotable than positive accomplishments, when the positive accomplishments, to date at least, are often only interesting descriptions of promising approaches. Although some systems have been shown to have validity or stability under some conditions, none has yet been demonstrated to have tremendous practical utility for psychiatry in the immediate future. On the other hand, he will find in the volume a very great deal of worthwhile discussion concerning the necessary attributes of a good classification system and the approaches necessary to demonstrate the utility and validity of any given system.

It is, of course, difficult to summarize a conference of this length and complexity in simple terms. However, a longitudinal survey of the material of the conference does appear to be in order.

The conference began with a discussion of Dr. Temkin's paper on the history of classification in the medical sciences. Three major historical approaches to classification were stressed:

1. The clinical approach which classifies diseases on the basis of their response to therapeutic intervention;
2. A classification system which focuses on the phenomenology of the illness; and
3. The approach which focuses on the etiology of the illness.

In his discussion of the paper, Dr. Brill pointed out that the names used to describe phenomena in psychiatry are a good deal more stable over time, and more meaningful, than are the various temporary systems for classifying psychiatric disease. Individual diagnostic entities have been juggled around in one way or another for many years, while the names describing indi-

vidual phenomena have stayed remarkably stable. Dr. Brill commented on the necessity for having both an empirical and a scientific classification which might be quite different and might even, side by side, be utilized for different purposes. Dr. Lehmann commented that diagnosis means through-and-through knowledge of a disease and doubted that we have this kind of knowledge in most psychiatric conditions.

Dr. Zubin offered a definition of disease as "an ongoing process which, unless attended to, leads to shortening of life, as distinct from a defect which simply leads to a severe reduction of efficiency." (Later in the conference someone suggested humorously that military service would qualify as a disease under Dr. Zubin's definition.) At that point Dr. Pasamanack offered, as an operational definition, that disease was something that physicians deal with.

A general issue which reappeared again and again throughout the conference was first raised by Dr. Brill in describing the diagnosis as the most important single piece of information for clinical administrative purposes in the care and treatment of patients in a mental hospital. He recognized clearly that whenever one classifies a patient by giving a diagnosis, or assigning him to a typology for that matter, one loses information. A recurring question throughout the conference was whether the advantages to be gained for research or administrative purposes from giving a patient a discrete diagnosis, or assigning him to a specific subtype of, for example, schizophrenia (whether the subtype be based on old-fashioned clinical diagnosis, or on a new-fangled typology), was more useful for either clinical or research purposes than providing a more detailed description of the patient. In short, do the virtues of categorizing people outweigh the resulting loss of information? Clearly, for certain epidemiological and administrative purposes, the assignment of patients into

categories is useful. For other purposes, such as predicting outcome of treatment, there was less unanimity. Some of the newer typological approaches to the classification of patients into various subtypes had been shown to have some relevance to treatment response or treatment outcome.

On the other hand, the possibility remained, as stressed by Dr. Gleser for one, that for certain purposes such as the prediction of clinical response one might do better to take into account knowledge of a number of aspects of the patient's history and psychopathology as individual predictors of outcome. One might therefore better use a multivariate (e.g., many-variable) approach to the prediction of therapeutic response or clinical outcome, omitting the step of classifying the patient into a given disease entity or diagnostic subtype.

Intermediate positions were suggested by a number of participants including Dr. Torgerson from the statistical side, and Drs. Klein and Cole from the clinical side. Certain patient characteristics—for example, degree of excitement or inhibition, duration of illness or marital status—might be generally predictive of outcome, independent of the symptom constellation or psychiatric disease entity in which these characteristics were present. Other patient characteristics might be more powerfully related to outcome if considered as a constellation of phenomena leading to a classification or diagnosis.

The two arguments and the intermediate positions were quite clearly presented and constitute a major problem for this conference and a major task for future psychiatric research.¹ Can the field of psychiatry advance more rapidly and more effectively if the characteristics of patients

¹ These arguments have to do with the function of classification in prognosis in clinical psychiatry; the role of classification in science, in general, and psychopathology in particular is not, however, confined to problems of prediction. This issue is considered later on in this chapter and from a slightly different perspective in the final chapter.

are considered as interrelated in a complex manner most meaningfully handled as diagnostic entities which can then be studied in a greater depth? Or should the attributes of psychiatric patients be considered as relatively independent variables which interact in a complex manner but which can best be studied individually or in their interrelationships with other variables without undergoing the information loss inherent in assigning patients to diagnostic entities?

It became clear during the conference that positions favoring classification or diagnosis, as against the study of patient characteristics without classification or diagnosis, were by no means confined to the participants representing psychiatry. It was true, in general, that the psychoanalysts found grave difficulties in assigning patients to specific diagnostic categories and preferred to describe the presenting symptoms, while hospital psychiatrists felt that the assigning of diagnosis, if carefully done, had real clinical meaning and was a necessary prerequisite to good clinical practice. Even at that end, the position was by no means that diagnosis in its present form dictated treatment but that the study of the individual case necessary to determine and arrive at a diagnosis brought with it an increase in the clarity with which the case was understood, and this study, plus the diagnosis, suggested certain directions for treatment and management.

The greatest interest during the first day revolved around a suggestion originally made by Dr. Gardner that psychiatry adopt a classification system resembling that used by the American Heart Association. Dr. Gardner had had considerable experience studying the diagnoses utilized by psychiatrists working in the community and in clinics in the Rochester, N.Y., area, as these were reported to the Monroe County case register. From his examination of his data the psychiatric diagnosis assigned seemed to play a minimal role in the operational clinical judgments that the psychiatrists

were called upon to make. Only a small proportion of the diagnoses in the APA nomenclature was used at all and a handful of those terms seemed to be used as waste-basket diagnoses. His recommendation was that an expanded diagnosis should include the specific reaction of disorder indicated by the present classification system, but should also describe the phase of the illness, the level and course of the disability, the patient's basic personality and adjustment pattern, and perhaps the patient's response to some specific therapy. Dr. Gardner also noted that as psychiatry moved out into the community, situations were more and more frequently encountered in which the proper clinical focus was not necessarily the individual being brought to treatment, but the pathological family setting, or the social situation as a whole. This point was later reinforced by Dr. Jenkins and others who pointed out that in the case of certain criminals and juvenile delinquents, the subject himself was normal, but the social situation from which he came was pathological. On the negative side, Dr. Weinstock noted that he had discussed the Heart Association system with an expert in that area, and had been told, "This system is all very nice, but it's not good enough for the expert, and it's too good for the inexpert."

There followed considerable argument, unresolvable at this meeting, as to the possible utility of this approach in psychiatry. Although optimism was expressed about this approach, some argued that one had enough difficulty getting clinicians to use a single one-step diagnostic procedure without asking them to carry out a more complex diagnostic procedure which would involve five or six components. Others felt that clinicians would be more comfortable with a four- to six-point diagnostic system than they would when faced with making a single one-term diagnosis.

There was also a good deal of argument, particularly on the first day, about the number of psychiatric diseases which could coexist in a single patient. Apparently it

is possible, though difficult with the present nomenclature, to assign two or more psychiatric diagnoses to a given patient. Logically there seemed no reason to assume that the presence of a single psychiatric entity precluded the presence of others. On the other hand, as Dr. Hamilton pointed out, a patient might have any number of medical conditions at one time, but he usually presented himself for treatment for only one of them.

There was also a good deal of discussion as to what constituted the desirable or necessary raw material on the basis of which patients should or could be classified. In this area the papers presented at the conference covered a broad range. At one extreme in the study of Drs. Gerard and Mattson, carried out at Ypsilanti State Hospital under the auspices of the University of Michigan, a variety of biochemical, physiological, and psychological test measures were utilized in an attempt to empirically separate schizophrenia into major subtypes. At the other extreme were studies like those of Drs. Overall and Hollister which utilize the Behavior Profile Rating Scale, with its 18 items succinctly describing various aspects of psychopathology observed during a psychiatric interview, as the sole and sufficient basis for classifying patients into types. Social history data were utilized in a major way only by Dr. Garnezy in his research on reactive and/or process schizophrenia, although Dr. Katz' typology was based on the description of patients provided by relatives, concerning their behavior immediately prior to hospitalization. Types based on descriptions by relatives were shown to resemble closely the patterns of the clinical ratings assigned such patients by psychiatrists and to have a certain amount of predictive validity where treatment was concerned, thereby stressing an important fact in clinical research—relatives' descriptions of the patients' disturbed social behavior represent another source of significant and valid information in the investigation of psychopathology.

During the discussions presented by Dr. Hammond for personality theory and Dr. Clausen for sociology, it was clear that in the opinion of these experts, at least, neither personality theory nor sociology offered any immediate solutions to the problem of classification in psychiatry. Although appropriate theoretical approaches growing out of experimental psychology, followed by appropriate testing of the theory in clinical and/or laboratory situations, might eventually help clarify the problems of psychiatric classification, Dr. Hammond expressed little optimism that such an outcome was to be anticipated in the immediate future. Dr. Clausen noted that sociology, in recent years, has essentially abandoned the classification of societies, social situations, or even roles in any rigid manner and has concerned itself more with the major dimensions of either societies or roles.

Dr. Kety, speaking for biology, began by stating that classification was an essential part of the process by which science came to grips with the universe. He noted that Linneaus had developed a detailed, complex, and careful classification of plants into species with a strong belief that the species was a fixed phenomenon. However, relatively rigid classificatory philosophy did not prevent Darwin, who began as a taxonomist and classifier, from developing important and imaginative concepts about the changes in species occurring during the processes of evolution. Similarly, another example of biological classification, pharmacology, has been relatively successful using a multidimensional classification system in which the same drug may be classified both as a tranquilizer and as an antiemetic.

In general he felt that classification, in and of itself, was essential as an ordering of the phenomena in the world. Secondly, although such classifications have to begin by being frankly and blatantly phenomenological, they have to carry with them firm standards of what constitute hard phenomena, and primary data, as against inferences and hypotheses. Thirdly, a

classification system should not permit inferences which assume or avoid important questions which a classification system should help solve; e.g., a classification system which assumes that schizophrenia is either genetic or reactive is settling an unsolved question by fiat and is discouraging necessary research. Fourthly, a classification system should be able to tolerate and entertain uncertainty.

In the first half of the second day, attention was given to statistical procedures for classifying or typing individuals. It became clear that there were sound statistical methods for determining whether two groups of individuals were, in fact, different from each other (discrimination) and there were excellent and reliable statistical procedures for the assigning of individuals to one or another group within a classification system once the characteristics of each group were solidly defined. On the other hand, statistical methods for determining whether a heterogeneous group of individuals described on a number of quantifiable characteristics contain 1, 2, or 17 clearly separable diagnostic entities did not currently exist. Dr. Barnard, who had been invited to attend this meeting but was unfortunately unable to be present, was, however, believed to have made some progress toward developing a method for determining whether a heterogeneous group of individuals who had been described on a single measure included representatives of several populations.

Dr. Greenhouse clearly drew the distinction between probabilistic approaches utilizing statistical methods and nonprobabilistic approaches. Probabilistic approaches allow one to say with certainty the likelihood with which the given results could have been attained by chance alone. For example, in a discriminant function analysis to determine whether a group of neurotics is different than a group of schizophrenics across a number of measures, a test of significance may reveal, for example, that there is only one chance in a thousand that these

groups could be, in fact, the same when they appear to be different.

On the other hand, the majority of techniques currently being used in this field to cluster individuals are nonprobabilistic, as is the oldest of these approaches, factor analysis, itself. One may obtain clusters of individuals by intercorrelating the profiles of test scores of large groups of individuals (correlations between persons) and then subjecting this data (the intercorrelation matrix) to some form of cluster analysis.² One cannot, however, determine whether the clusters which are obtained might not have occurred by chance. For this reason, Dr. Greenhouse characterized such statistical approaches as descriptive. They might be extremely valuable in identifying potentially valid clusters of individuals, "patient types," but they could only be considered the first step in the development of hard scientific fact. A classification or a typology derived from such descriptive statistical approaches must be followed up by experimental tests of its validity; i.e., relating the system to other criteria which establish the likelihood that the typology actually exists or to criteria which indicate that the typology has the potential for predicting other phenomena, other behavior, e.g., response to treatment. He noted that statisticians were a good deal more comfortable with the clustering of a number of measures—for example, items in a rating scale—into factors by factor analysis than they were with the clustering of individuals into factors, by the same method.

Dr. Lyerly pointed out that several conceptually different approaches were being used by investigators employing statistical methods to subdivide a population of psychiatric patients into groups. Some investi-

² This example of a descriptive clustering procedure obviously does not do justice to the variety of such procedures which have been used or the technical complexities of the actual operations involved. The controversy about the appropriate agreement score, and the bases for selecting among several clustering procedures for a particular problem, are topics which were taken up in depth in the "Issues of Methodology and Statistics" section of the conference.

gators used indices of agreement such as correlation, while others used measures of dissimilarity such as distance as a starting point. The particular statistical approaches they selected were in part based on the purpose of the classification. If it is necessary that all patients be classified, e.g., for hospital administrative purposes or because the investigator wishes a total system, then it may be more appropriate to start by assuming that all patients are in one large cluster, which is then subdivided successively into smaller clusters. Where the intent is to determine how many natural clusters exist in what appears to be a heterogeneous population, then the investigator may start with an individual patient as the basis of a cluster to which others who are most similar to that patient are added. The latter approach usually results in a number of small clusters with many individuals left unassigned to any cluster, while the former approach guarantees that every individual will be assigned to some cluster.

Dr. Rao discussed statistical methods for assigning individuals to groups, given specific criteria for inclusion or exclusion from such groups. He stressed the importance of having statistical methods—essentially decisionmaking rules which would not only permit the assignment of an individual to predetermined groups, but also the determination that a given individual observed did not belong to any of these groups. Only by such a system can one discover new groups.

There was a good deal of discussion following this, considering the problem of the normal distribution. A good many statistical procedures assume normal distributions of the measures included in the analysis. The example most familiar to psychiatrists is the normal distribution found with intelligence test scores where the mean is 100 and the standard deviation is 15; a bell-shaped curve results if the IQ's of a number of individuals in a large random sample of the population are plotted on a graph. A good many variables observed in psychiatric patients are not normally dis-

tributed, even within a population of psychiatric patients, much less throughout the total human population. For example, in a population of acute schizophrenics the majority will not show visual hallucinations, while a small proportion will. This produces a J-shaped distribution on a rating of degree or intensity of visual hallucinations for the group as a whole.

A similar problem of obvious importance both to statisticians and psychiatrists was the question of linear relationships. Statistical techniques most commonly utilized in research on classification in psychiatry assume a straight line relationship. For example, the more anxiety one has, the worse one's performance will be on a psychological test. On the other hand, clinical experience suggests that for many variables, including anxiety, the relationship between two variables may be better described by a U-shaped function. For example, with very, very low anxiety, performance on a psychological test may be impaired through inattention or apathy. With a moderate level of anxiety performance may be optimal, while at a very high level of anxiety, performance may again be impaired through disorganization of coping behavior. To the extent that such phenomena, yielding nonlinear relationships, occur to any large extent in psychiatric conditions, the more conventional statistical approaches may fail to find significant relationships which do in fact exist. These problems are basic to the application of statistical procedures in the field of classification, but it was clear that although they must be recognized they are not necessarily insurmountable. Many of the statistical procedures used are sufficiently robust to overcome partial violations of the assumptions of normality and linearity and to result in accurate conclusions. It is clear, however, that where these data indicate gross deviations from these conditions, other procedures, which are designed to reflect the actual nature of such deviations; e.g., measures of curvilinear relationships,

are more appropriate and should be applied.

The longitudinal problem, the fact that a particular disorder is characterized by several phases over time, is also raised. It seems reasonable that the same clinical condition, at different points in time, might look relatively different. In that case, an empirical classification system which is based only on a cross-sectional view of the patient might fail to detect the similarity between an acute paranoid schizophrenic in an acute exacerbation of his illness, and essentially the same kind of patient when he was beginning to enter into a period of remission.

During the afternoon and evening sessions papers describing actual studies of psychiatric patient groups were presented. The discussion then broadened to consider the kinds of data concerning patients which psychiatrists or other trained observers can assess with reliability and precision. Dr. Grinker, for example, noted that in his study of the characteristics of depressed patients, the patients' feelings and concerns were rated with considerably greater reliability than were their ward behaviors.

Dr. Lorr presented his approach to the cluster analytic problem and typology of psychosis based on an analysis of the psychiatric symptom data from large representative samples of hospitalized patients. Essentially similar subtypes had emerged from the analysis in five different bodies of data. Samples of psychotic patients had been drawn both from admissions to Veterans Administration hospitals and from those admitted to non-VA installations.

Dr. Overall presented a summary of the results of a continuing program of research which has been utilizing a variety of statistical methods, to develop typologies which are then examined for their ability to predict differential treatment response. A series of interesting findings have emerged, in studies both of schizophrenic and depressed patients—for example, anx-

ious depressions do better on thioridazine (Mellaril) while retarded depressions do better on imipramine (Tofranil)—but this work is based on relatively modest sample sizes with some groups being compared not being treated concurrently. This work requires replication.

Dr. Katz reported on a series of studies in which a typology (which is based on relatives' ratings of the social behavior of acute schizophrenics prior to hospitalization) was empirically derived by a modified version of Q analysis. He distinguished between two kinds of validity, consensual and predictive, and presented evidence for both. It was found that psychiatrists' independent ratings of the same patients using a different scale confirmed the patterns of behaviors which were identified by the relatives. Secondly, the typology was found to be predictive of response to treatment; an "acute panic" state, for example, was found to be more responsive to phenothiazine treatment than other types in two separate studies.

Dr. Zubin's paper emphasized the need for new methods in the biometric approach, as against the clinical approach, to diagnosis. He noted that many possible tests suffer from being culture—or even social class—related or dependent on motivation and suggested some newer objective methods which were relatively free of these limitations. He also speculated on the possible independence of pre-illness personality and illness-caused psychopathology. If these were independent, difficulties in utilizing certain types of measures as diagnostic devices would be inevitable.

Drs. Saunders and Gittinger described a new method for utilizing factor analytic studies of Wechsler-Bellevue Intelligence Test items to yield a much larger variety of subscores from this instrument. These scores would then permit its use not only as a test of current personality functioning but as a reflection of past problems and future trends. Evidence for the validity of this procedure was not presented and thus

the technique obviously requires further investigation.

Dr. Stein discussed some of the problems and limitations inherent in developing typologies based on self-descriptions; including the observation that workers developing dimensions of personality or typologies may restrict their own results by the limits of their sample selection—often to college students willing to volunteer for personality testing.

On the third day Dr. Garmezy discussed his comprehensive and excellent review of the historical development and research data relevant to the concepts of process and reactive schizophrenia. His general conclusions were that the process-reactive distinction reflected a continuous clinical spectrum rather than two clear and separable clinical entities and, second, that the general concept was experimentally valuable and productive of new knowledge. He suggested also that a similar continuum probably existed in other psychiatric conditions.

Dr. Klein, in his work, also hypothesized a continuum—a bipolar dimension of agitation—excitation versus retardation and withdrawal which was drug responsive at its extremes but not near its midpoint. This dimension coexisted with other types of psychopathology; if severe at either extreme, diagnosis was difficult. When either excitement or retardation had been resolved, the presence or absence of other pathology would determine the overall clinical outcome. He also stressed the prevalence of nonlinear relationships between psychiatric phenomena.

Dr. Clyde's unique utilization of canonical correlation demonstrated that two groups of patients existed who differed in the way their response to drugs was assessed by doctors and nurses. In one group, the doctors and nurses agreed about the degree and direction of clinical change; in the other they disagreed. The pretreatment characteristics of both groups were brought

out by this complex probabilistic technique, infrequently applied to psychiatric data.

Dr. Gerard then reported on a massive study which utilized the best available clinical, historical, psychological, biochemical, and physiological measures to describe a large group of schizophrenic and nonschizophrenic patients hospitalized in a public mental institution—an excellent summary of the hitherto unpublished results of a massive study. In addition to yielding knowledge about differences between schizophrenics and nonschizophrenics and reinforcing the importance of the paranoid-nonparanoid dichotomy within schizophrenia, the study also yielded seven schizophrenic subtypes based on a Q-analysis of psychological, physiological, and biochemical measures. This typology was validated, in considerable part, when these groups were found to differ meaningfully in their clinical history and current psychopathology.

Dr. Fink presented a sound review of past efforts to use the EEG, either resting or activated, as a basis for diagnosis or typing. His report reflected both some past inadequacies in research in this area and its current promise.

The conference finished with summary discussions by Drs. Katz, Cole, Hamilton, Zubin, and Rao. In the present volume, the two former discussions have been replaced by a more considered final chapter by Drs. Katz and Cole analyzing the results and possible long-range implications of the conference proceedings.

To do a capsule summary in this introductory chapter, the following points can be stressed:

1. The current APA nomenclature has a number of defects which make it clumsy to use for a variety of clinical and research purposes. Serious consideration might be given to expanding it to include more kinds of information, as is done in the American Heart Association system.

2. A refreshing variety of research approaches to the classification of psychiatric patients has been seriously employed in the past few years. These approaches vary in the type of data used as a basis for classification, in the statistical model used for developing groups and in the extent and nature of the validation work accomplished to date.

3. Serious and detailed consideration has been given to the many and varied purposes classification can serve in psychiatry, as well as its role in related disciplines. It was clear that its differing functions in the practice of psychiatry and in the science of psychopathology influence both the phenomena which are selected for study in these spheres and the statistical methods which are used for developing a new system. This issue was considered in some depth at the conference and will be explored in further detail in the final chapter.

4. Somewhat unexpected attention has been given to the history of classification in

psychiatry with emphasis on the underlying models used.

5. The values and limitations of classifications or typing of patients versus the consideration of a variety of patient characteristics for their relevance to research or clinical management without formal classification have been discussed from a variety of viewpoints.

6. The ideal characteristics of a classification system and the research steps necessary to establish the value of a new system are considered in detail.

In short, this volume gives a unique overview of the current state of research in classification in psychiatry and psychopharmacology and a valuable coverage of the possible roles of classification, its values and limitations, its history and philosophy. Although the volume does not immediately solve all the problems it raises, it is designed to provide a sound footing for better future research in this important area.

The History of Classification in the Medical Sciences

Owsei Temkin, M.D.¹

1

I must begin my presentation by pleading ignorance. My topic is the history of classification in the medical sciences (1). Yet I do not know as of when exactly we may speak of medical science, leave alone medical sciences, and when classifications are first to be found in them. But I feel confident that classification existed in medicine long before it became scientific in any sense of this word, and that in late antiquity at least, classifications of medicine existed which showed a remarkable viability.

In ancient Egypt, a primitive specialization of practitioners went together with a gross division of diseases (2). There were those who treated diseases of the eyes, others who took care of diseases of the head, still others of those of the teeth or of the abdomen. This division was reflected in a specialized literature, and one such monograph is available in the Edwin Smith Papyrus (3), which deals with surgical injuries as they might be suffered in ancient warfare. The author apparently followed an arrangement from head to foot, though actually the text stops in the middle of the 48th case, where the upper thorax has been reached. Making an anachronistic use of logical categories, we can say that a genus of diseases, viz external injuries, was classified according to the anatomical seat.

Somewhat similarly, in ancient Mesopotamia, series of medical texts were arranged according to different complaints: child-

birth, afflictions of the lungs, of the eyes, etc. Religious and magic elements also played a role. Labartu attacked children and pregnant women, and Nergal brought on pestilence. What exactly led to all these distinctions is hard to tell, yet it seems important that the origins of medical classification go back into a remote past, and that the beginning of the medical sciences lies somewhere along the road, not at its start.

There existed at least one archaic type of classification that entailed a conscious effort, viz. enumeration. The ancient Egyptian physicians knew of the relationship of heart-beat and pulse, and they speculated on the number of vessels that led to various parts of the body carrying humors or air and causing diseases. The so-called Knidian books of the Hippocratic collection favored enumerations of diseases: Three kinds of phthisis, four diseases of the kidney, and so on.

Numerical schematism could be imposed by an inherited predilection for certain numbers, and it could be combined with numerical speculations. Empedocles fixed the number of elements at four, and these four elements with their contrasting qualities of hot-dry, hot-wet, cold-dry and cold-wet were accepted by Aristotle and remained the basis of chemistry until the Renaissance. In medicine, the Hippocratic works accounted for health and disease by various humors; one book of the collection, however, listed bile, phlegm, blood, and black bile (4). What exactly led to the distinction of these

¹ Director, Johns Hopkins Institute of the History of Medicine, Baltimore, Md.

four humors is not known. If we give free rein to our imagination we may think of various explanations connected with the color of these humors or the stratification of blood, or with other observable data. Or we may say that the number four related the humors to the four elements and thus to the universe. At any rate, the fact remains that those four humors became classical and that correspondences were seen between the four elements, the four qualities, the four humors, the four ages of man, and the four seasons of the year. Thus macrocosm and microcosm were comprised in one harmonious picture.

Moreover, the four humors served to classify human temperaments. In late antiquity continuing into the middle ages, four psychosomatic types were differentiated, each according to the predominant humor which also determined psychological makeup (5). There came into existence the sanguine, the phlegmatic, the choleric, and the melancholic temperament, still remembered today, though usually divested of the somatic characteristics. The scheme is an example of the confluence of philosophical and physiological speculation, observation, and numerical bias, where it is next to impossible to segregate the individual role of each factor. The example also illustrates the ambivalent nature of all classification: on the one hand, to name, number, and sort what has struck man as different and, on the other, to search for order among what appears to be amorphous.

Yet we have already moved into a time when classifications of medical sciences did exist. Two schemes, closely related, confront us at the transition from antiquity to the middle ages (6). Both divided medicine into theory and practice. The first scheme numbered five major disciplines. Of these, physiology, pathology, and symptomatology belonged to the theoretical branch, whereas hygiene and therapeutics constituted practical medicine. The other scheme tells us more about the things to be studied. Avicenna, the great philosopher-physician who

lived about A.D. 1000, gave it a poetical form. "Medicine," he said, "is the preservation of health and the cure of disease that has made its appearance from a bodily cause. Its first division is into theory and practice" (7). But from here on this classification deviated from the first named. Expressed in the barbaric terminology of the middle ages, the theory of medicine was supposed to encompass seven natural things, six nonnaturals, and three things contrary to nature. The seven natural things were those of which the human body is composed and formed, and by means of which it functions, *viz.* the elements, the temperaments, the humors, the solid parts, the faculties, the functions, and the spirits. The six non-naturals were food and drink, air and other external, as well as internal, factors that determined whether the organism was in health or disease. Finally, the three contranaturals were disease, its symptoms, and its causes.

The ancient classifications not only prevailed into the 17th century but were still taught in medical schools and textbooks of the 18th century under the name of "Institutes of Medicine." The "Institutes" set forth the principles of medicine under five headings: Physiology, pathology, semeiotics (symptomatology), hygiene, and therapeutics, obviously the five disciplines with which we started, though the material content had changed considerably. The scheme dissolved around the end of the 18th century, when anatomy and chemistry emancipated themselves, and when physiology took on the more restricted meaning it has today.

The scheme presented the framework for some of the classifications of Galenic medicine, which dominated medicine down to the 17th century. According to Galen, health consisted in man's feeling well and his ability to exercise the functions of daily life. Galenic medicine, therefore, needed a classification oriented toward function. However, function was the work of anatomical structures endowed with or helped by various faculties, natural, vital, or psychic.

With Aristotle, Galen distinguished between similar parts like bone, vein, artery, nerve, fat, which were believed to be structurally homogeneous, and organic parts which were constructed of a variety of similar parts. The similar parts were composed of the four elements and their qualities. When their proper balance was disturbed, the similar parts were distempered, and this constituted primary diseases. These diseases were rooted in the physico-chemical state of the tissues, if the use of such modern terms for the sake of clarification is allowed. Organic diseases were those in which shape or size or composition or position was abnormal. Severance of continuity belonged to both classes (8). In the 16th century, pathological classification had not changed much (9).

This system of pathological classification was adapted to Galenic pharmacodynamics. Like everything else, drugs too had their qualitative formula, and a very hot primary disease would be treated by a correspondingly cold remedy. The characteristic qualities of both diseases and drugs were, of course, integral parts of Galenic science; they made sense only within that system. However, Galen also provided two different classifications of drugs. The first, according to the diseased part, ran from head to foot, beginning with remedies for alopecia and ending with prescriptions for arthritis and podagra (10). The second grouped remedies according to their therapeutic effects, e.g. those that made wounds cicatrize, those having an emollient effect, etc. (11). That meant a much stronger intrusion of the practical and traditional way of classifying disease than was the case in his pharmacodynamic classification.

2

Long before the rise of medical sciences, the Greeks as other nations before them, had names for diseases. Hippocrates tells us that "the acute diseases are those to which the ancients have given the names of

pleurisy, pneumonia, phrenitis, and ardent fever, and such as are akin to these, the fever of which is on the whole continuous" (12). The names for diseases increased considerably in course of time, and later it became fashionable to write works on acute and chronic diseases, a division represented by the textbooks of Aretaeus and Caelius Aurelianus.

We cannot accommodate all ancient diseases in our modern nomenclature. Lientery and kausos are strangers to us. Nor do we feel at home with explanations offered in terms of ancient pathology. Yet the method of describing a disease by clinical symptoms or syndromes is familiar to us. We may call this procedure empirical. Indeed, the ancient medical sect of Empiricists, who cultivated the observation of diseases and of remedies on the basis of clinical observation without recourse to science, used pathognomonic syndromes for the definition of diseases (13).

The empirical approach never disappeared from practice, and it assumed added importance towards the end of the 17th century. Galenic science had broken down under the impact of the new mechanics and chemistry. But these new sciences were as yet unable to serve the medical practitioners as reliable guides. Sydenham and Baglivi stressed the need for unprejudiced histories of disease, and thus there came into existence the classical nosological systems of the 18th century, represented by the names of Boissier de Sauvages, Cullen, and Pinel (14).

Nosology is the classification of disease, and in the 18th century it appeared as the practitioner's science. It was not meant to abolish the other sciences, but it was to offer what they lacked: immediate pragmatic value.

What was the physician to do when anatomy, experimental physics, and mathematics are not yet able to offer guidance, Boissier de Sauvages asked. He should turn to Sydenham, who had formulated the rules for nosology; species and genera of diseases

ought to be distinguished on a purely factual basis in contrast to causal theories, which should be kept apart, and the character of a disease should be derived from its constant symptoms. Thus a disease was to be defined by the enumeration of the symptoms which suffice to recognize its genus or species and to distinguish it from others (15).

On this basis, Boissier de Sauvages divided diseases into 10 classes, each with its orders and genera (16). His system of symptomatological nosology was bolstered by his criticism of anatomical and etiological classifications. In view of the fact that our standard nomenclature of diseases and operations is built on these two principles, the criticism is of particular interest. Anatomical classification, according to the parts of the body as the seat of disease, Boissier de Sauvages declared confused, inconvenient, and deceptive. All parts were interconnected, they did not exist in isolation, and they pointed neither to the cause of the disease nor to its treatment. Of many diseases we did not know the seat at all. "What is the seat of mania, melancholia, forgetfulness, sleepwalking, tarantism, vertigo, catalepsy, nightmare, pica, and a thousand others?" (17)

True, Boissier de Sauvages admitted, the human body was a machine, and the same causes must produce the same diseases. But a cause, *qua* cause, is nothing sensible. Causes only render the possibility of diseases understandable, not their actual presence. The symptoms were the sensible factor. From the symptoms we could go back to the causes; but neither the seat of the disease nor its cause leads to the symptoms.

Like Boissier de Sauvages, Cullen, too, pleaded for the practical value of nosology. Nosology was to be a textbook of pathognomonics (18). Consequently, Cullen eliminated such symptoms of a disease as showed themselves over a lengthy course of time. Causes might be admitted as legitimate characters; but only if they were

sufficiently certain and easily to be observed (19). The result was astonishing. Syphilis, for instance, belonged to the Class: Cachexiae, order impetigo, and it was defined as "Contagious; ulcers of the tonsils appear after impure venery and disease of the genitals; clustered pimples appear on the skin, chiefly at the margin of the hair, going off in crusts or scabby ulcers; pains in the bones, and protuberances of some parts of them" (20). I have taken the example of this disease with its changing manifestations over the years to illustrate the insufficiency of this pathognomonic classification.

With Pinel, the last of the great nosologists of the 18th century, we enter the world of Condillac and his disciples, the French ideologues, with their faith in the method of analysis. This method led Pinel to distinguish six primitive febrile orders on the basis of least equivocal symptoms (21). Anatomical changes were included, although Pinel was not always clear about the relationship of these fevers to postmortem findings. But in the class of phlegmasias, which, by definition, were caused by inflammatory conditions, he not only insisted on an anatomical principle but, criticizing Morgagni, shifted the emphasis from organs to tissues.

Not the simple position of the parts, but the correspondence of organic structure and vital function must serve as a guide. Thus the phlegmasias will be divided into different orders according to whether they have their seat in the mucous membrane, the diaphanous membranes, the glands, the muscles, and the teguments. For instance, what difference does it make that the dura mater, the pleura, the peritoneum are located in different parts? Must they not be brought together into the same order when they show analogous lesions in the state of phlegmasia? Their particular differences establish the generic character, and thus the acute diseases offer a whole that is just as exact as the fevers and even more complete because of the light thrown by the phenomena of post

mortem dissection upon the picture of symptoms (22).

As is well known, this stimulated Bichat to his research on tissues, whereby he became the founder of modern histology. It is, moreover, not difficult to find passages in Pinel that link him with Bichat's demand to combine "rigorous observation with an examination of the changes which our organs show" (23). Bichat, in turn, exerted a great influence upon Laennec, with whom the anatomical trend of the Paris school gained the upper hand (24).

3

Pinel's observations and classification owed much to the opportunity the Paris hospitals offered for observing many cases of the same disease at different stages, so that disease could be visualized as a process possibly extending over years. It could also be approached statistically, just as the effect of drugs could be statistically evaluated. Last, but not least, the hospital facilitated the pursuit of pathological anatomy by combining in one place clinical wards and dissecting rooms.

To the influence of the hospital, there was now added the influence of the surgeon, who had risen to equality with the physician both socially and scientifically. To the surgeon disease was not a mere natural entity comparable to a plant or an animal. Injuries were visible and their cause usually known, and even in tumors, cataracts, hernias, abscesses, where the cause might not be known, the seat at least was certain and all-important not only for diagnosis, but for operative treatment as well. As we have it from Laennec and others, the new anatomical-clinical diagnosis tried to extend the surgical point of view to internal diseases (25). If possible, their anatomical and clinical symptoms were coordinated, and percussion and auscultation helped to do this even during the patient's lifetime. Gastric ulcer became a disease entity; Louis described typhoid fever with its character-

istic inflammation of Peyer's patches; Bright, in London, associated morbid changes in the kidney, edema, and albumen into an entity that came to be named after him. These examples will suffice to indicate the trend toward a classification of disease according to organs. The trend furthered specialization, and specialization, in turn, furthered the trend. Specialization in the 19th century was largely a matter of increased knowledge about certain organs and of more complicated methods of diagnosing and treating their abnormalities. Internal disease and external disorders had previously demarcated the domains of physician and surgeon, with obstetrics coming in as a belated third partner. In course of the 19th century, ophthalmology, urology, dermatology, otorhinolaryngology, neurology, cardiology, gastroenterology came to denote not only specialties but also a classification of diseases according to their organic seat.

4

Anatomy was an old science. Over the span of more than 2,000 years, it had developed a classificatory system that was geared to the anatomical part as a bearer of function. But between 1800 and 1858 anatomy, normal and pathological, underwent a decisive reform which resulted in a new classification. Bichat replaced the old "similar parts" by 21 tissues, morphologically different, and also different in their reaction to stimuli. About 40 years afterwards, Schwann offered a cellular theory that explained the development of tissues from cells. Then Henle synthesized this theory in his textbook of "General Anatomy." The body consisted of unformed elements, i.e. chemical compounds and mixtures, and formed elements, the cells, which built the tissues. Physiological properties, for instance the contractility of muscular tissue, were still included in this anatomical work (26).

Pathological anatomy followed the development of normal anatomy step by step. But it was not enough to say that

there were morbid organs, morbid tissues, and morbid cells. On each level notions had to be formed and classified.

Gross pathology had borrowed many of its concepts from surgery. Inflammation, necrosis, gangrene, atrophy, ulcer, tumor, abscess, all these existed on, in, and below the skin. Much of this classification could be applied to tissues, but this was not enough. For instance, there was the question whether new kinds of tissues evolved in pathological processes. Laennec differentiated between formations that were analogous to existing ones and those without analogy (27). Similar questions arose with the shift from tissues to cells. Virchow based his cellular pathology on normal cellular histology. But he had to explain the appearance of disease in the cells morphologically as well as physiologically.

Cells, Virchow taught, were accessible to three types of irritation: functional, nutritional, and formative. Nutritional irritation might lead to simple hypertrophy, but it might also lead to degenerative changes marked as cloudy swelling. Formative changes might lead not only to a multiplication of cells, hyperplasia, but also to the formation of changed cells, metaplasia. Classification and neologisms went hand in hand, because this classification was not a mere regrouping of known things but a classification of material not observed before.

5

A look at Austin Flint's "Treatise on the Principles and Practice of Medicine" allows an insight into the changes effected in nosology about 100 years ago. The part of the book devoted to special pathology arranged diseases on organic principles very much as we find it in textbooks today, except for the last section devoted to fevers and other general diseases. In this area, the anatomical approach admittedly had broken down, but a new one had not yet been found. Hence many older nosological notions survived. There existed primary,

idiopathic, or essential fever, one of the elementary forms of disease, (28) a survival of Pinel's idea. This kind of fever was different from symptomatic fever that may accompany other diseases. Flint classified fevers into "Febricula" which he identified with Cullen's synocha or inflammatory fever, (29) and the well-known types of continued, periodical, and eruptive (exanthematous) fevers with their subdivisions.

It needs no reminder that this province was to receive its reform from the science of bacteriology. Seen from the point of medicine, bacteriology had its obvious place. It elucidated so-called infectious diseases by establishing their causes. It took quite a time to realize that the existence of the specific pathogen was a necessary but not the only condition for the specific infection. In the meantime, bacteriology effected a more rigorous distinction between syndromes, clinical entities, and diseases, the latter name being reserved for entities whose cause had been established. Bacteriology thus revived nosology which, as we shall see, was in danger of succumbing to the onslaught of the physiologists.

Although the medical significance of bacteriology is uncontested, bacteriology per se is not a medical science. From the 18th century to the second half of the 19th, bacteriology proper was in the hands of botanists who tried to classify bacteria. One of the sponsors of Robert Koch's early work was the botanist Ferdinand Cohn. Yet we usually think of Koch rather than Cohn as the founder of bacteriological methodology. Only when the medical significance of bacteria became clear, did bacteriology become a significant branch of biology. As I understand it, a classification of bacteria and, "a fortiori," of viruses is hardly possible without reference to the conditions of life of the microorganism in the host organism, including man. The lesions which the *Trepanoma pallidum* causes are part of its characteristic behavior, as are the ill effects it suffers from such substances as arsphenamine and penicillin. From the bacteriological point of

view that is probably all that is necessary. To say that this micro-organism causes the disease syphilis is to translate this behavior into medical terms.

6

But I have anticipated the course of events, for a hundred years ago, bacteriology had not yet been generally accepted by medicine. Even the claim of cellular pathology to represent general pathology was still very young. Nevertheless, this claim already encountered competition on the part of experimental physiology, which refused to be a corollary of anatomy.

Physiology was the science of the phenomena of life. Claude Bernard had a grand vision of the classification of these phenomena into the two groups of organic destruction and organic creation (30). And general pathology, as Julius Cohnheim treated it, was the pathology of those phenomena which were the object of study of normal physiology. Thus his lectures began with the pathology of circulation and of nutrition (31).

Claude Bernard denied the existence of fundamentally different medical sciences. In the final analysis, there was only one medical science, experimental physiology under its two aspects, normal and pathological. It was a causal science; it taught the factors that preserved health or brought about disease. It also included the scientific part of therapeutics, for it taught how various substances acted upon the body and how they influenced pathological conditions. This was the proper subject of experimental pharmacology, which Buchheim considered a part of physiology. Thereby an entirely new orientation was given to the study of drugs, traditionally comprised by the name of *materia medica*.

Cullen had classified drugs according to their physiological or therapeutic effects. There were astringents, tonics, emollients, corrosives, stimulants, sedatives, refrigerants, antispasmodics, and so on. Around the mid-

dle of the 19th century, *materia medica* was a huge and uncritical collection of drugs, Galenic as well as chemical, which practitioners used with more or less confidence, a few abstaining from drug therapy altogether. There was no lack of classifications. The textbook by Jonathan Pereira distinguished between empirical and rational arrangements, the latter having five divisions, which were in turn further subdivided (32).

But the new pharmacology left it to the physician to select what he might find useful. Schmiedeberg, Buchheim's pupil, followed his teacher in rejecting a classification of drugs according to chemical principles or to action on important organs. The pharmacologist ought to follow the botanist and pay attention to all such attributes of the effective agents as were pharmacologically important. "Those substances," Schmiedeberg said, "whose properties and effects agree most with one another * * * are united in pharmacological groups, and each of them is named after one of the best known substances among them" (33). The result would be "a natural system" which could be perfected with the advance of scientific pharmacology (34).

All so-called medical sciences, then, could be considered parts of physiology. The question was only whether there remained anything in medicine that was not a mere practical application of physiology. In particular, what was the status of nosology, the classification of disease?

Radical physiologists repudiated all classification of diseases as entities. At best these were practical makeshifts, having neither real existence nor scientific claim. For disease was a modification in the working of the organism that had to be explained, not labeled by arbitrary names of alleged entities.

Hughlings Jackson's attack of 1866 against the nosological concept of epilepsy as a disease entity is illustrative. "The word 'epilepsy,'" he suggested, "should be degraded, and be used to imply the condi-

tion of nerve tissue in sudden and temporary loss of its function, whether that be loss of sight, loss of consciousness, or 'running down of tension' in those parts which govern muscles" (35). Instead of speaking of genuine epilepsy, chorea, etc., the student should always try to establish the functional changes in the nervous system that accounted for the patient's symptoms.

But about 10 years later, Jackson admitted the necessity for both a scientific and an empirical classification, side by side. A botanist, he explained, classified plants scientifically; the gardener did so from a utilitarian point of view. Likewise, medicine "qua" science needed a classification that advanced knowledge, whereas medicine as an art needed a clinical classification that was practical, even if arbitrary (36).

7

Jackson's last remarks underline a feature which, I trust, has become visible in course of our sketch. For instance, we noted that scientific pharmacology left it to the therapist to classify and to select the drugs that he needed for curing diseases. Somewhere in the background was always reference to the classification of the diseases which medicine was to prevent or to cure. What then is the significance of nosology, the existence of which all medical sciences seem to presuppose? If it is nothing but a makeshift, why then does it loom so large? (37)

The pertinence of this question can be brought home by a look at the history of classification of psychoses during the last century (38). Around 1800, insanity was the object of psychiatric classification. There was little doubt that insanity manifested itself in various forms. Pinel distinguished mania, melancholia, dementia, and idiocy, to which Esquirol added partial insanity or monomania. Both Pinel and Esquirol believed, however, that some kind of transformation between these classes was possible (39). The unitarians, Guislain, Zeller, and Griesinger went so far as to claim that

the various forms were nothing but stages of one morbid process.

With the rise of anatomical investigations, a pathological anatomy of the psychoses was fervently looked for. But the quest was not very successful, except in the case of general paralysis, for which there emerged a coordinated anatomical and clinical picture. This success had far-reaching consequences for Kraepelin's eventual classification of mental diseases. Kraepelin united Kahlbaum's catatonia, Hecker's hebephrenia, and Morel's "démence précoce" into a definite entity, *Dementia praecox*. Kraepelin's other famous nosological creation was manic-depressive insanity (Falret's "folie circulaire"). These and other mental diseases were the object of definite clinical diagnosis, on which Kraepelin insisted. Diagnosis was so important because it gave the correct basis for prognosis. Kraepelin spoke of "the certainty with which our disease concepts allow us to predict the future march of events" (40). He did not and could not know the infectious origin of general paralysis. Otherwise he might have classed it with other infections and might have hesitated to pattern the concept of other mental diseases after it.

Different as the schemes from Pinel to Kraepelin were, they helped communication, they allowed statistical comparisons, and they reflected matters of great concern. The diagnosis of insanity often meant segregation of the patient, and legal questions could be connected with the diagnosis of monomania. The future of an individual was predicted by general paralysis or *dementia praecox*. The nosological classification was the map which the psychiatrist used in discerning between health and disease; it led to prognosis and to an indication of therapy or its uselessness.

You may remember the example I offered of the bacteriologist who scientifically described the action of *Treponema pallidum* on the host. Nothing compelled him to mention the disease, syphilis, and to conjure up its diagnostic, prognostic, and social as-

sociations. But without such associations, would the biology of this pathogen have aroused the interest it received? Would Ehrlich and Hata have synthesized arspenamine, and would this drug have created the pharmacological sensation it evoked, if it had not promised a cure of that disease?

The history of classification in the medical sciences has been shown to be different for every one of them. It was seen to depend on at least two factors: the development of the science to which it referred, and the development of medicine. By medicine I here mean the discernment between what is healthy and what is disease. This discernment (diagnosis) is dictated not only by science but also by what the individual and society consider normal and abnormal. This discernment also classifies diseases so that their future course can be gauged and an indication for treatment obtained. Such classification has more connotations than a medical dictionary offers. For diseases are acute or chronic, dangerous or slight, curable or incurable; they require expensive treatment or not, they are shameful or pitiful, serious or ridiculous, and they bear whatever other signatures life, custom, and history impress upon them.

Only when the medical categories of health and disease and the conceptual tools for grasping them, viz diagnosis, prognosis, indication, and therapy are brought to bear upon a science does it acquire the nature of a medical science. And diagnosis, by definition, includes classification of the things to be discerned. This seems to be the moral taught by the history of classification in the medical sciences, and this may also explain why classification in medicine is older than the medical sciences are.

REFERENCES

(1) Because of the ramifications of this paper, I cannot cite all the literature that has a bearing on the topic. General acknowledgment is made to Knud Faber, "Nosography: The Evolution of Clinical Medicine in Modern Times," second edition, New

York: Hoeber, 1930. Richard Koch, "Die ärztliche Diagnose," Wiesbaden: Bergmann, 1917. Ralph L. Engel, Jr., "Medical diagnosis: present, past, and future," *Archives of Internal Medicine*, 1963, 112: 512-519, 520-529, 530-543. Walther Riese, "The Conception of Disease: Its History, its Versions, and its Nature," New York: Philosophical Library, 1953. Further literature on related subjects will be found in Owsei Temkin, "The scientific approach to disease: specific entity and individual sickness" in A. C. Crombie (ed.) "Scientific Change," New York: Basic Books, Inc., 1963, pp. 629-647. "Historical aspects of the classification of science" is the title of an appendix (A) in B. C. Vickery, "Classification and Indexing in Science," second edition, New York: Academic Press, Inc., 1959.

(2) For ancient Egypt and Mesopotamia see Henry E. Sigerist, "A History of Medicine," vol. 1: "Primitive and Archaic Medicine," New York: Oxford University Press, 1951.

(3) James Henry Breasted, "The Edwin Smith Surgical Papyrus," Chicago, 1930 (University of Chicago Oriental Institute publications, vols. III and IV).

(4) "On the Nature of Man."

(5) The history of the four temperaments has been investigated by Raymond Klibansky, Erwin Panofsky, and Fritz Saxl, "Saturn and Melancholy," New York: Basic Books, 1964, pp. 55 ff. and 97 ff.

(6) Owsei Temkin, "Studies on late Alexandrian medicine I," *Bulletin of the History of Medicine*, 1935, 3: 405-430. *Idem*, "Byzantine medicine: tradition and empiricism," *Dumbarton Oaks Papers*, No. 16, 1962, pp 95-115.

(7) Avicenna, "Poème de la médecine," ed. Henry Jahier et Abdelkader Nouredine, Paris: Les Belles Lettres, 1956, p. 12.

(8) This simplified sketch is based on Galen's book "On the Distinction of Diseases," Kühn edition, vol. 6, p. 836 ff.

(9) See Ioannes Fernelius, "Universa medicina," *Pathologia* 1, c. 6; Paris: Wechel, 1567, p. 180.

(10) Galen, "On the Composition of Medicaments according to Regions," Kühn edition vol. 12, p. 378 ff.

(11) Galen, "On the Composition of Medicaments by Classes," Kühn edition, vol. 13, p. 362 ff.

(12) Hippocrates, "Regimen in Acute Diseases." Translation by Jones, Loeb Classical Library, vol. 2, p. 67.

(13) Temkin, "The scientific approach to disease," op. cit. (above, note 1), p. 632.

(14) On nosology in the 18th century, see Lester S. King, "The Medical World of the Eighteenth Century," University of Chicago Press, 1958, ch. VII, p. 193 ff.

(15) Franc. Boissier de Sauvages, "Nosologia methodica," ed. C. F. Daniel, t. 1, Leipzig: Schwickert, 1790, Prolegomena, article 74; p. 33.

(16) A convenient survey of the classification can be found in Gulielmus Cullen (ed.) "Synopsis nosologiae methodicae," editio sexta, Edinburgh: Gulielmus Creech, 1803. This volume, apart from several nosological systems not discussed in this paper, also contains Cullen's own nosology in Latin.

(17) Boissier de Sauvages, *op. cit.*, Prolegomena, 62; p. 25.

(18) William Cullen, "Nosology" (anonymous English translation), Edinburgh: Bell and Bradfute, 1810, p. v.

(19) *Ibid.* p. xviii.

(20) *Ibid.* p. 150.

(21) Ph. Pinel, "Nosographie philosophique, ou la méthode de l'analyse appliquée à la médecine," 2 vols., Paris, an VI; vol. 1, p. xiv.

(22) *Ibid.* p. xxv.

(23) Xav. Bichat, "Anatomie générale," vol. 1, 1920 Paris ed. p. 50.

(24) According to Erwin H. Ackerknecht, "Laennec und sein Vorlesungsmanuskript von 1822," *Gesnerus*, 1964, 21: 143 n., 8, Laennec made a veritable cult of Bichat.

(25) Owsei Temkin, "The role of surgery in the rise of modern medical thought," *Bull. Hist. Med.*, 1951, 25: 248-259.

(26) J. Henle, "Allgemeine Anatomie, Lehre von den Mischungs- und Formbestandtheilen des menschlichen Körpers," Leipzig: Voss, 1841.

(27) R. Th. Laennec, "Note sur l'anatomie pathologique," *Journal de médecine, chirurgie, pharmacie, etc.*, an XIII, 9: 360-378; see also Ackerknecht, *op. cit.* (above, note 24).

(28) Austin Flint, "A Treatise on the Principles and Practice of Medicine," Philadelphia: Lea, 1866, pp. 109-693. Although Flint considers microscopic research, the book shows the influence of the Paris school rather than of contemporary German medicine.

(29) *Ibid.* p. 697.

(30) Claude Bernard, "Leçons sur les phénomènes de la vie communs aux animaux et aux végétaux," vol. 1, Paris: Baillière, 1878, p. 125 ff.

(31) Julius Cohnheim, "Lectures on General Pathology," Translated by Alexander B. McKee, 3 vols., London: The New Sydenham Society, 1889-90.

(32) Jonathan Pereira, "The Elements of Materia Medica and Therapeutics," 2 vols., second American ed., Philadelphia: Lea and Blanchard, 1846.

(33) Oswald Schmiedeberg, "Grundriss der Arzneimittellehre," zweite Auflage, Leipzig: Vogel, 1888, p. 12.

(34) *Ibid.* p. 13.

(35) Quoted from Owsei Temkin, "The Falling Sickness," Baltimore: Johns Hopkins Press, 1945, p. 308.

(36) See *ibid.* p. 310.

(37) I am here speaking of nosology not as a mere schematic table of names and definitions of diseases, which has been rightly criticized by W. Bieganski, "Medizinische Logik," Würzburg: Kabitzsch, 1909, p. 75 ff., but as representing the knowledge and associations connected with the names of diseases.

(38) Erwin H. Ackerknecht, "A Short History of Psychiatry," translated by Sulammith Wolff, New York: Hafner, 1959; Norman Dain, "Concepts of Insanity in the United States, 1789-1865," New Brunswick: Rutgers University Press, 1964; Gregory Zilboorg, "A History of Medical Psychology," New York: Norton, 1941.

(39) For Pinel see Evelyn A. Woods and Eric T. Carlson, "The Psychiatry of Philippe Pinel," *Bull. Hist. Med.*, 1961, 35: 14-25 (see pp. 19-20).

(40) Preface to the fifth edition of Emil Kraepelin, "Psychiatrie," Leipzig: Barth, 1896, p. v (Quotes correspond to the printing of the German text).

Discussion

LEON EPSTEIN, M.D., *Associate Medical Director, Langley Porter Neuropsychiatric Institute, San Francisco, Calif.*,
Chairman

(Paper: "The History of Classification in the Medical Sciences,"¹ Owsei Temkin, M.D., Director, Johns Hopkins Institute of the History of Medicine.)

OPEN DISCUSSION

(Dr. Epstein outlined Dr. Temkin's paper and served as moderator of discussion from the floor.)

Dr. BRILL. The historical perspective is a very important one in classification of psychiatric disorders, and it is a pleasure to comment on this very scholarly report on this aspect of the problem—especially because Dr. Temkin agrees with some of the ideas I happen to have, one of which is the need for diagnosis and classification in psychiatry and apropos of this I would like to read his concluding paragraph:

Only when the medical categories of health and disease and the conceptual

¹ Advance text received and distributed; author not present at conference.

tools for grasping them; namely, diagnosis, prognosis, indication, and therapy are brought to bear upon a science does it acquire the nature of a medical science. Yet diagnosis by definition includes classification of the things to be discerned. This seems to be the moral taught by the history of classification in the medical sciences and this may also explain why classification in medicine is older than the medical sciences are.

One can say bluntly we cannot have medicine without names, and no psychiatry without names. We can't even have ordinary communication without names; our daily language consists of a series of names, which are nothing more than classifications of individual objects into groups such as men, iron, rain, weather, etc. Common usage is that we call the process "naming" when we classify individuals into groups, and we use the term "classification" when we arrange groups into a system of higher order. Thus for example, we name substances as iron or copper, or cats, dogs, or horses, and such a name collects individuals into a group. In medicine and in psychiatry diagnosis can be seen as a name of an entity using "name" in this sense.

An example of classification in our sense is seen in the Linnean system which gathers the names into family, genus, and species in a hierarchical pattern. Generally such a system is based on some overall field theory of some combination of unifying concepts. At its best this type of unifying theory gives us what we have in Darwin's theory of evolution or in the Mendelian periodic table. At its worst, we get the kinds of field theories which sprang up in the late 1700's and early 1800's and are described so vividly in Dr. Temkin's paper. We should remember that any field or unifying theory is likely to look a little bit silly to the people who come afterwards and we should have enough insight to realize that whatever theory we use today for our grand strategy of classification may in 20 or 30 years look quite as ridiculous as some of the old

theories we now look down on. This will probably depend on basic sciences.

The basic units of classification or what we called names of specific syndromes are far more durable than overall systems. If you read the ancient writers such as Coelius Aurelianus and medievals like Paracelsus you will find depression described quite accurately, alcoholic psychoses quite accurately, epilepsy with its mental complications very well, and so forth. These entities or concepts were not created *de novo*, but were the result of a long evolutionary process which gradually brought them to their current level. The classifications on the other hand, that is the gathering of these names into systems, have had a very different story. These have been juggled around and rearranged again and again. I recognize that along the road—and Dr. Temkin pointed this out—various workers have reversed the process and created totally new systems and have also provided new names to fill out these systems. Here, history is very instructive, too. These totally synthetic systems which were relatively independent of what went before have generally lasted about as long as their originators. It's extremely difficult to create a system *de novo*. Maybe now with new techniques we can do it, but we should recognize that this is not easy.

The intent of this conference is compatible with the historical perspectives and principles brought out in Dr. Temkin's paper; it is not to do away with psychiatric classification but rather to seek an answer to the question, what is the best classification or set of classifications to serve our needs.

With respect to the repudiation of disease entities I would like to cite Dr. Temkin's paper. He says that Hughlings Jackson, in 1866, attacked the nosological concept of epilepsy as a disease entity. In addition to scientific reasons Hughlings Jackson had a close female relative who had epilepsy, and it is possible he didn't like the term epilepsy because of the bias it involved, and he might

have been tempted to solve the problem by saying: "There is no such thing as epilepsy."

However, as Dr. Temkin points out: "About ten years later Jackson admitted the necessity for both a scientific and an empirical classification side by side. A botanist, he explained, classified plants scientifically; a gardener did so from the utilitarian point of view. Likewise, medicine as a science needed a classification that advanced knowledge, whereas medicine as an art needed a clinical classification that was practical even if arbitrary," and this in one paragraph is a summary of the paper I have written for today.

Jaspers in his "General Psychopathology" also covers this particular point; namely, that the form of classification depends on its purpose; there are many purposes and thus many classifications of any group or any class of material. I would like to add that no classification is to my mind, no science, and no medicine.

Dr. LEHMANN. The chain of factors, basic factors, that Dr. Temkin pointed out are, after all, crucial, namely: diagnosis, prognosis, indication, and therapy. It sounds easy and fundamental.

On the other hand, I think there is a hidden measure of trouble, particularly with two of these concepts: diagnosis and prognosis. "Indication" is a rather simple signal for some operation, and "therapy" is also operationally defined. "Diagnosis" and "prognosis" are not.

"Diagnosis" is easy when there is a case of measles or chickenpox. One knows the incubation time; one knows what caused it, and how things will develop; one by one the symptoms can be predicted.

But what do we mean by "diagnosis" in other diseases? Just the symptoms or also the cause? What is going to happen naturally, or how a patient is going to react to treatment? Every one of us probably has his own private concept or personal equation of what he means by diagnosis.

One gets into similar difficulties when one proposes the concept of an endogenous depression. Again and again, with other psychiatrists, students, and residents, you run into the same argument. How do you know there is such a thing as an endogenous depression? You just don't know what the traumatic loss was which caused the depression. The loss is perhaps subthreshold. It's a reaction to a severe loss that occurred in early childhood, and it is a symbolic loss that the person suffered yesterday, perhaps, when getting a letter he didn't like which reactivated the deep trauma that occurred when he was 3 years old. Therefore, what you are really dealing with is a reactive depression. You just don't know the trauma; therefore, you call it endogenous. But all depressions are really reactive. So their argument runs.

Well, if we had an operational definition of "diagnosis" we could say that any depression not associated with a readily recognizable trauma within a few months prior to its onset is called endogenous. That would secure our diagnosis against the allegation that the depression may be a symbolic reactivation of an early loss, and so on.

I think we have to look seriously into the question of operational definitions in diagnosis—and the same is true for prognosis. What do we mean by prognosis? Disappearance of symptoms, rehabilitation occupationwise, or social rehabilitation? That the patient will be as well as before, or better, or what? We lack operational definitions of what we mean by prognosis. The difficulty, as I mentioned, is compounded—or rather concealed—by the fact that indications for therapy are by their very nature operationally defined. Because of this we simply take it for granted that we also know what we are talking about when we say "diagnosis" and "prognosis."

Dr. HAMILTON. I think one of the most important sections of Dr. Temkin's scholarly paper is where he described the situation in general medicine; i.e., problems of classi-

fication in a field which we now understand. It is a great pity he did not expand this particular section, because the whole point of this paper is that the history of classification of disease in medicine is a subject from which we can learn a great deal. He pointed out we are now precisely in the position of Flint, 100 years ago, when he was trying to define the fevers and describe and classify them. He had a special category of primary, idiopathic, or essential fever, and you may have noticed that one of the papers at this conference is concerned with this type of approach, reactive versus endogenous or primary. The most important lesson we have to learn from the history of classification is how other people have tackled problems similar to ours.

The points made by Dr. Lehmann are sound and cannot be repeated too often. What I was trying to say about classifying bruises, headaches, and rheumatism together, on the basis of their response to aspirin, is not that it could not be done but that it was a trivial classification.

The purposes of classification, as Dr. Brill has just said, are manifold. There are many reasons why we do it, and we should appreciate that some of the reasons may be important and others may be trivial.

Now, in the case of the "aspirin classification," it is important only insofar as the physician gets to the point of writing down on a sheet of paper the prescription for aspirin. From the point of view of the prescription, yes, the three disorders are alike, and treatment isn't trivial, but when it comes to prognosis, causation, and etiology, they are quite different, and these differences are extremely important. There are many ways of making classifications and from the operational view they are equally good, but there is a hierarchy of importance, and that is the point I am trying to make.

Dr. KLERMAN. I am sorry Dr. Temkin is not here, since I feel hesitant about expressing disappointment with a paper whose author is not present. I'm not sure whether

my disappointment is with Dr. Temkin's paper or with the people who prepared the program. I feel that either there is a chapter missing from Dr. Temkin's paper or that there is a paper missing from the program. No matter which is the case, I feel that a gap has been left in the subject from the age of Kraepelin until today.

As I understand it, Dr. Temkin's point—and this brings me to my view of our current dilemma—is that the principle of classification, which developed in the 19th century as a result of the introduction of bacteriology and other biological sciences, has been very successful in most fields of medicine and is embodied in the standard nomenclature and the international code. Applied to psychiatry, however, it has produced great dissatisfaction in most of us. The etiological approach seems to have bogged down for the majority of patients with whom we come in contact in our everyday clinical life.

Confronted with this breakdown, we can adopt either of two positions: We can say that we just do not know and will have to wait until etiological knowledge catches up with practice, or, as seems to be the case in the research of the past decade, other points of view may be brought in to classify these patients according to principles which do not derive from the 19th century model of classification by etiology.

I had hoped that Dr. Brill, as chairman of the APA committee,² might respond to this. In the past 20 years the APA nomenclature seems to have reflected the dilemma of mixed classificatory criteria. The first section, on acute and chronic brain syndromes, uses an etiologic classification derived from the 19th century. Then the paper moves into a very mixed type of classificatory principles, as Dr. Gardner points out in his paper, sometimes based on primarily behavioral categories, sometimes based on presumed etiologic factors.

² Dr. Brill later explained that he had been elected to APA Council, and his position as Chairman of the Nomenclature and Statistics Committee was held by Dr. Gruenberg.

In conclusion it seems to me, either a chapter is missing from Dr. Temkin's paper, or another paper is missing from the program which would bring us a scholarly condensation of the transition from the grand tradition of the 19th century to our current dilemma.

Several conference members suggested a fundamental question is: What do we mean by disease?

Dr. ZUBIN. Disease is usually defined³ in terms of "aetiology," "structure," and "symptomatology." In most mental disorders aetiology is unknown, the structure of the organs of the patient (as far as we now know) unaffected, and symptomatology is the only available basis for the definition. Mental diseases whose aetiology and structural defects become known, are usually lost from psychopathology. Thus disorders like general paresis, pellagra with psychosis, epilepsy, even PKU are now largely in the hands of other disciplines. Only the diseases of unknown origin remain in the field of psychopathology. Furthermore, there is also the question of whether mental disorders are in fact diseases or merely reaction patterns. In order to bypass these philosophical dilemmas and still have a working definition, we shall define mental disorders (and at the same time any kind of a disorder, mental or physical) as follows: Mental disorders are of two kinds: diseases and defects. A disease is a progressive condition which, unless attended to will result in premature death, extreme reduction of efficiency and happiness (severe inner distress), or both of these. A defect is a more or less stationary condition which leads to severe reduction of efficiency but not necessarily to premature death.

[Dr. LEHMANN pointed out that in German psychiatry you cannot claim compensation for someone who has developed some

sort of a neurosis, let's say an obsessive-compulsive one, because it is not considered a disease but a personality disorder. Here, that would mean we would have to throw out most of our psychoneuroses from the disease classification and put them under something else.]

Dr. PASAMANICK. Our classifications are manmade and sociologically derived. They have been constructed for use.

This has already been said and I don't think we can reemphasize this too often because we keep approaching the elephant from different sides, blinded, of course, and then say the other man is wrong.

The present medical classification serves many functions at the same time; unfortunately this must be so at this time because the physician serves many functions at the same time.

It is also incomplete and will be incomplete as long as we have gaps in knowledge. Therefore, we have what seem like inconsistent classifications. We have a part of a classification etiologically derived because it serves a certain function, both in terms of research on these conditions and in terms of dealing with them at the same time, as well as other purposes. Another portion is phenomenologic, a third may be anatomic, and a fourth histologic, but all subserve different functions.

All of these functions are served simultaneously and this is one of the troubles we have with the classification. Nevertheless, I think we have to live with these functions because they serve certain uses.

The definition of disease Dr. Zubin presents is, within its limitations, a very adequate one. But let me point out there are other definitions of disease which are socially derived.

For instance, under the definition Dr. Zubin presents there are conditions such as delinquency, poverty, or others which can fall into that classification or that definition if you use it strictly. But it is not what the physician deals with directly because that

³ Definition amplified for publication.

hasn't been given to him as a social responsibility.

We must constantly remember where these classifications come from historically and what their present usage is. Medicine is a growing body of knowledge and experience. By looking at the past, we can see that as new techniques and new knowledge are acquired, the classification systems that are

favorable are preferred in a given generation; some change. As Dr. Brill pointed out, some varieties of classification used today are going to look silly 20 or 30 years from now. But I think there is a more general proposition in Dr. Temkin's paper: We hope all varieties of classifications we use today will look somewhat foolish 30 years from now in the light of later knowledge.

THE VARYING ROLES OF CLASSIFICATION IN PSYCHIATRY AND PSYCHOPATHOLOGY

The purposes of classification are viewed from the standpoint of the clinician, the epidemiologist, and the researcher in psychopathology. Representatives of the biological and social sciences comment on its meaning and functions in their respective fields. There is discussion of the values classification serves in the clinical and in the basic sciences, its varying roles in description and prediction in psychiatry, and its role as an essential methodological step in the development of the science of psychopathology.

The Role of Classification in Hospital Psychiatry

*Henry Brill, M.D.*¹

Introduction: The Special Place of Classification in the Operation of Mental Hospitals

If mental hospital practices are in any way an expression of functional needs one must assume that classification of patients by mental disorder must serve a significant set of needs because such classification is virtually universal and a routine practice in all mental hospitals and the practice is not new; in fact one of the earliest steps in the organization of mental hospitals in their modern form was the creation of such a system by Pinel who wrote in his "Treatise on Insanity" (1806): "The term 'mental derangement' is sufficiently well adapted to express the various lesions of the human understanding. But to obtain accurate knowledge of this subject we must not remain satisfied with general expressions however suitable and comprehensive. For this purpose we must examine the different species of derangement analytically, consider them separately and thence deduce the principles both of medical treatment and hospital management."

The needs of mental hospitals long continued to dominate the field and further development of psychiatric classification remained for well over a century in the hands of mental hospital psychiatrists of whom the best known are probably Griesinger, Kahlbaum, Kraepelin, and Bleuler. It is only relatively recently that extramural psychiatry has emerged as a strong influence

and the accommodation of its special needs has resulted in important changes which are still in progress.

The close relationship between the history of the mental hospital and that of psychiatric classification has had a curious form of general recognition and that is a fusion between stereotyped concepts of both. According to this notion, classification in the mental hospital, and especially in the State hospital type of facility, consists of an elaborate but futile and degenerated form of intellectual exercise, a sterile professional gesture which replaces and blocks constructive therapeutic effort. As is often the case with serious mistakes, this one has some factual basis; it refers not to the present, but to a time some thirty years ago when the methods and ideas of Bleuler and Kraepelin had begun to lose their newness but were being carried forward in a routine fashion still encumbered with all the niceties of differentiation developed when they were research tools. This propensity was undoubtedly intensified by the fact that the therapeutic drive of medical men had in a sense to be sublimated into observation because of the severe limitations of treatment techniques available at that time. This emphasis on description had its counterpart in physical medicine too and also reached a level of virtuosity which the young physician of today can no longer emulate nor does he wish to do so. In both cases one can ask how much was contributed by the great effort which was channeled into pure description but I feel that there is no

¹ Director, Pilgrim State Hospital, West Brentwood, N.Y. Formerly Chairman, American Psychiatric Association Committee on Nomenclature and Statistics (1960-65).

more reason to think that it actually impeded the advance of psychiatry than of medicine generally. There can be no doubt that for a time it was an end in itself in the hands of the unimaginative or the uninterested and that it was misused but today under a grinding pressure for professional time, classification in mental hospitals has long since reverted to a level which corresponds roughly to that of physical diagnosis (note I did not say laboratory or X-ray) in a modern general hospital and if psychiatric classification is still universally employed in mental hospitals one must seek some more fundamental reason than the dead hand of tradition. Classification does have the powerful sanction of professional accrediting bodies but this is to be seen as recognition of its value and not as the basic cause for its existence.

Much can be and has been said in criticism of existing systems and particularly that of the current (1952) APA Manual and the validity of these criticisms remains a continual spur to change and to improvement but to avoid the tendency to throw the baby out with the bath water in our enthusiasm we should identify also the positive and constructive side of nosology in mental hospitals. This will provide us with information as to what values we should seek in any system which is to replace the current one. From this point of view the place of classification in mental hospitals will be described under several headings.

Internal Administration of the Mental Hospital

The patient's case record is the key to any kind of team effort in medicine and to any type of large-scale medicine, and the mental hospital represents both. In the case record, which serves as the corporate memory of the organization, the classification or diagnosis represents the ultimate condensation or abbreviation of medical information, and in its brevity lie both its strength and its weakness. It is a law of communication that

reliability varies inversely with abbreviation and that redundancy lends security and certainty. Thus it is obviously a fallacy to hope that a single phrase can ever communicate the significance of an entire case record and this had often been stressed by those who oppose the use of such labeling. On the other hand, with all of its penumbra of uncertainty, the classification is a useful point of departure for any review of a case record and the initial orientation conveyed by such a term as "chronic brain syndrome associated with alcohol" is quite different from that of "anxiety reaction." In an active hospital service where one must make a multitude of decisions of all types with respect to patients whom one may know casually or not at all, case records are under continual review and because the diagnosis can be seen at a glance it is regularly one of the first items to be checked. Of course this is only a start of a record review, the inquiry is then broadened until there is enough information for a decision to be made and personal contact with the patient is often required.

The value of the diagnosis in any given case will depend on the skill, experience, and quality of work which lies behind it but these factors are available for evaluation as one goes over the case and in this respect also evaluation of psychiatric diagnosis is comparable to that of diagnosis of any other type. Hasty diagnoses hastily arrived at can be valueless or worse and in this fact lies the great danger of the current tendency to downgrade diagnosis as unnecessary, redundant, and even harmful; this attitude has all the elements of a self-fulfilling prophecy since diagnoses produced in such climate of opinion are likely to be all that has been said against them and it would appear that for the protection of the patient it is one of the continuing tasks of hospital administration to maintain adequate standards in this area. Later, when we consider the place of classification in training of staff we shall have an opportunity to see how diagnosis can provide

leads for the monitoring, documentation, and treatment of patients.

External Administrative Relationships

Classification of patients is as important for external communications as it is within the hospital but here we find some of the major difficulties which have led to a condemnation of all psychiatric diagnosis. A large and growing list of agencies have an interest in diagnosis as related to such issues as disability, mental competence, and criminal responsibility. They even try to find quantitative measures in categories which were never intended to measure degree or intensity of syndrome and a confusion between diagnosis and prognosis continually reappears in various guises. All of this leaves us with a reluctance to reveal the formal classification and this is intensified by the suspicion that the labels carry a degree of stigma and may be used against the interests of the patients since some informed laymen even in administrative positions still look upon psychiatric patients as dangerous, and either unrecoverable or subject to recovery that is labile and unreliable.

On the other hand, these same terms of classification are accepted as proof of disability and open the way for various types of special consideration which patients often require; the words "schizophrenic reaction" reveal the lazy vagrant to be a sick man and the goldbricking soldier looks different when he is seen as suffering from a psychoneurotic or a psychophysiologic reaction and after a diagnosis, even the aggressive drunkard and the narcotic addict are found more to be treated than punished. Thus for many of our patients the psychiatric diagnosis confers a social benefit though for a vast majority the stigma outweighs all advantages. For this reason, even those of us who recognize the value of diagnosis in some contexts will wish to withhold it at other times and present instead a phraseology which will convey the facts more correctly though less briefly and less specifically.

In the last analysis the possible misuse of a valuable tool is hardly sufficient grounds for abandoning it and it would seem far more logical to continue instead with the process of public education and so to establish the correct significance and the limitations of the terms which we use.

At the present time it is a practical fact that there are many situations where the formal diagnosis is demanded and the welfare of the patient requires that it should be provided. Judges, insurance carriers, and various agencies in the health and welfare field wish explicit statements on which to base their actions and here every hospital has to maintain a nice balance between its responsibilities to its patients, past and present, and toward society in general; it seems hardly likely that at the present time any public hospital will be able to evade the issue by refusing to make diagnoses altogether.

Classification as an Aid to Teaching

The average mental hospital is in a perpetual process of renewing its staff, and for the psychiatrist in training classification of the mental disorders of patients has a special value.

If properly carried out, classification is not a simple and single act of applying a label but represents the culmination of a complex process of medical logic based on adequate documentation which includes a good history, physical and mental examination, and well-chosen laboratory work, in addition to a series of adequate and well-oriented psychiatric examinations all based on careful, direct, and well-informed observation. The diagnosis is not and can never be final and absolute and the outcome of psychiatric logic is thus quite different than that of mathematical logic but it is analogous in at least one respect—in either case one can guess at an answer—and one may even guess right on occasion, but this is subject to supervisory check and the need to arrive at a conclusion is an important stim-

ulus for a complete well-rounded examination while the requirements for diagnosis are a valuable guide to the young physician in eliciting relevant facts and selecting decisive details from the bewildering mass of possible data which is the story of the average person with psychiatric illness; even the most extensive case record does not contain everything and diagnostic considerations are an important aid in picking out what is important. Another important characteristic of adequate classification is that it requires attention to the matter of differential diagnosis which is a further stimulus to thought and a powerful instrument of education. In addition it has the further advantage of counteracting any tendency to simplify psychiatry into one disorder with one treatment method.

Classification on this level is an exacting mental task and sometimes becomes rather tedious so it is not surprising that the young physician may be easily persuaded that diagnosis is outmoded, useless, and a waste of time, and time is a burning issue in the active services of the mental hospital where the case material is very mixed and the pressure of work is highly conducive to various rationalizations for reducing it. The fact that our current classification practices have survived such a highly ruthless competition for professional time has its own significance.

The physician in training may at times believe that regardless of what they may say the example of senior men provides proof that classification is unnecessary because such experienced psychiatrists may and often do develop personal variations of the classification system or may even give up any conscious and overt reference to any known classification, but when stripped of their detail such private systems almost always resolve themselves into a paraphrase of standard classification so formulated as to emphasize a special orientation as to etiology, pathology, and/or treatment. Descriptions of such freedom from convention are easily misunderstood and misin-

terpreted and can lead to serious confusion on the part of the trainee; this is a case where it is far better to have learned and forgotten than never to have known at all and abandonment of classification based on ignorance remains ignorance.

In addition to its significance in day-to-day operations, classification represents an indispensable approach to the literature of psychiatry since a very large proportion of this literature is structured in terms of the concepts of clinical classification which all of the important modern systems tend to have in common. This is important for the psychiatrist's initial education and essential for his subsequent capacity to maintain contact with new developments. In this connection it is an encouraging and interesting fact that major psychiatric classifications have shown a long-term trend toward convergence of concepts and of terminology with a tendency to conform to a common pattern whereas at one time variation was frequent and extreme and virtually every important hospital center had its own highly individual system.

The Place of Classification in Evaluation of Treatment and of Hospital Operations: Statistical Issues

It is perhaps significant that except for investigators who are interested specifically in problems of classification as such, those who work with mental hospital data tend to do so in terms of one of the standard systems and this includes formal research in areas such as epidemiology, treatment techniques, treatment results, and genetics as well as routine statistical compilations based on ordinary hospital operations. The serious variability which is inherent in such use of existing clinical classifications is well known and unavoidable although devices have been developed to help control it in special studies; in routine work, reliance is placed on the fact that variability is reduced when essentially the same observers are involved for the entire period under

examination and when such period does not extend beyond a year or two.

When one moves into the area of comparisons from one country to another and one period of psychiatric history to another, variability is so great that clinical classification of patients is not usually suitable for statistical purposes though case comparisons are possible and one can still within limits identify various entities in case material from other countries and other periods of history.

The variability which has been noted is beyond a doubt due mostly to uncontrollable variation of standards and practices of classification but it is also quite likely that some differences may be due to true variations in the nature of the case material itself as may be suspected from the fact that ascertainable characteristics of patient populations such as age do vary very definitely from one location to another and from one time in history to another. Among the useful facts about classification is the observation that the outcome of treatment with various modalities tends to correlate with classification but it does so only in a qualitative sort of way and the overlapping of diagnoses and methods is so great as to obscure greatly whatever relationship may exist. Those who are most interested in such investigations feel the limitations of our current subdivisions of case material very keenly and continue to use these systems only because of serious doubts as to which if any, system is likely to produce more consistent results. In addition they point out that by utilizing an accepted system it is possible to make oneself understood by all professionals and to find some standard of reference, no matter how vague, in the world literature.

Here again one is faced by the fact that classification alone is not decisive for prognosis, treatment, or for any other important judgment in the individual case and these limitations are compounded when one deals with groups of cases although the contrary assumption is often made because of the

falsely quantitative appearance of columns of figures. These are limitations and not a complete nullification of value; if one allows for the limitations, figures based on classification can still give useful data on control of hospital operations and of treatment, and will undoubtedly continue to be used in that way until something better is developed.

A Time for Change

So far in this paper we have seen that classification of patients by mental disorder is a universal practice in mental hospitals and that the history of the development of a modern nosology has been closely linked with that of the modern mental hospital. We have also reviewed some of the ways in which formal nosology plays a role in the treatment and care of patients, in training of staff, in administration, and in community-hospital relations. All of this information refers primarily to the APA nosology which completely dominates the scene in the United States and indeed has a far wider use than any other system in the world.

The information we have considered is presumptive evidence that prevailing systems and especially that of the APA do have practical value, but it does not by any means prove that an ideal has been approached or that there is even general satisfaction with what we now have, and the fact is that the great preponderance of professional sentiment of all types, both in hospitals and extramurally is that it is now time for a change. Partly this can be traced to the fact that the current APA nosology has not had a significant overhaul since 1952, and this is well beyond the 10-year span which is the life expectancy of medical classifications generally. The mere passage of time is of course not important in itself, but it is a convenient measure of the amount of change which has taken place since the last revision and this has provided the motives for a steadily growing general conviction that constructive innova-

tion in this field is both necessary and possible.

In the course of the development of this body of opinion much has been written and said about classification and a wide variety of issues have been raised. Some of these are familiar and open to direct examination on the grounds of clinical hospital experience, others are of a more fundamental nature and involve broader principles in scientific logic, systematics, epistemology, and semantics and these create a serious dilemma since they are difficult to discuss without extensive digressions into abstract fields, but if they are left unanswered they can leave an impression that they represent compelling and established proof drawn from a higher realm of thought.

Among such issues one may include the question whether we should attempt to classify at all, either on the ground that there is no such thing as mental disorder but only a false analogy to physical medicine, or on the ground that there is only one mental disorder and so it is still incapable of being classified. Closely related to these issues is the question as to whether we are classifying people, illnesses, phenomena, symptoms, or constellations of findings. If issues are to be raised on this level they should receive full discussion on the same level. Brief hit-and-run references to such questions can produce only confusion. Suffice it to say that on the philosophical level the balance of evidence is clearly in favor of having a classification, and there is nothing unique about the philosophical issues raised with respect to psychiatric classification. These are general in nature and belong mostly to the science of systematics where the values of purely artificial classifications are well established; so far as this writer has been able to ascertain after a considerable amount of exploration in this field the essential questions are direct and empirical ones, and we are quite safe in discussing such matters on their merits in a direct and practical fashion.

With regard to the debate as to which

is the best classification one can say only that this is a question without meaning until one specifies for what purpose. For the purposes of the mental hospital it has been apparent for some time that the APA clinical classification is due for revision and streamlining if only for the reason that it has outlived the normal 10-year span, but in addition it has been criticized on many grounds and most particularly that it is unwieldy, its terms are too long and that some of them need to be dropped while new ones need to be added. Beyond this it is apparent that we shall be called upon to classify patients as to disability types which are not listed in the current APA Manual. One of the most urgent is that of social and economic disability. Thus it seems that we shall probably require not one new classification system but several, a clinical one and others for various special purposes, and at this point it seems that we can do no more than identify and point out the qualities which our experience indicates are useful in a clinical classification for hospital use, and our experience includes only the clinical type.

Requirements of a Classification for Mental Hospital Use

The following is a tentative list of qualities which appear necessary for any clinical classification to be used in mental hospitals:

1. That it be as simple as is compatible with accuracy. The generations of trainees who do a large proportion of hospital work succeed each other with considerable rapidity, and the ease and speed with which they can grasp the principles and apply them will determine how well it will be accepted and how effectively applied.

2. That the system be clinical. Whatever other characteristics it may have, any system which is to be applied clinically must be able to group mental illness of patients in a clinically meaningful way.

3. It cannot be too much emphasized that the new system should be as compatible as possible with those which have preceded it

and those which are in use elsewhere in the world. International communication grows apace and hospital psychiatry is a worldwide activity. What happens in England has the utmost relevance to what we shall be doing in this country as can be seen from the recent experience. The channels of communication should be as wide and as easy as possible and this means that all possible international characteristics of the language of psychiatry should be preserved to the very greatest degree possible. To create a really complete hiatus between ourselves and the past, and even worse to create an awkward incompatibility between our concepts and those of the rest of the world would be a serious matter indeed.

In this connection it should be noted that during the last few weeks (December 1965) the World Health Organization announced the appearance of its latest revision of the "International List of Causes of Disease" and the section of this document which is devoted to mental disorders will deserve careful attention.

The mental hospitals have long ceased to be the sole users of psychiatric classification but they still constitute an important sector of psychiatric practice and it remains to be seen how their needs will be harmonized with those of all other groups in the field; it is to be hoped that the classification will be designed in such a way as to be as broadly acceptable as possible; to quote Stengel: "In this field there are few absolute truths and there is little room for last ditch conflicts about matters of doctrine." The value of any classification will depend among other things on how many workers accept it, understand it, and use it and for this it should be free of considerations other than those of descriptive accuracy and clinical utility.

Finally, it should also be pointed out that for routine hospital use, the assimilation of a new system of classification is a major undertaking extending over a period of years; it would seem that for anything which requires a broad acceptance in hospitals

across the country or even internationally, the burden of proof should be on the innovators to show that a real advance and improvement is involved and not yet another self-contained system.

If we attempt too many changes we shall certainly risk losing what uniformity of terminology we have and create useless confusion and misunderstanding in mental hospital activities. On the other hand, the hospitals have always been able to adjust to change and are even now in an active phase of readjusting their activities to broad new alignments in the public health and welfare fields and these changes will undoubtedly have to be reflected in some way in our systems of diagnosis and classification. One of the tasks for psychiatry in the immediate future is to see to it that the changes are constructive, practicable, and that they will take into consideration the special operating requirements of all various users, including the mental hospitals, and this will include the requirement that they be simple, brief, by no means doctrinaire, and that all change will be fully justified on professional grounds.

The Role of the Classification System in Outpatient Psychiatry¹

*Elmer A. Gardner, M.D.*²

It is basic to an intelligible language that we can assign to a single object a collective concept, and hence, communicate,

¹ Supported by grant MH-00381-07 from the National Institute of Mental Health. I want to thank Miss Mary Lynch, Dr. Harold Miles, and Dr. Sanford Meyerowitz for the assistance they have given me in the preparation of this paper, their thoughtful suggestions, and their critical comments.

² Associate Professor and Director, Community Mental Health Center, Department of Psychiatry, Temple University Health Science Center, Philadelphia, Pa. Formerly, Associate Professor and Director, Division of Preventive Psychiatry, Department of Psychiatry, University of Rochester School of Medicine and Dentistry, Rochester, N.Y.

via word or symbol, characteristics and relationships shared by innumerable different objects. George Simpson points out further, that the ordering of these perceptions (or operationally, of the objects perceived) is the province of science, and is a reflection of the order of nature.

The process of classification of objects systematically related to each other is essential to the scientific method, and surely central to medical science. In psychiatry, however, we have been plagued through the years by an inadequate system of classification, or one which truly fails to afford the terminology necessary to present a lucid diagnostic picture of the individual. Ralph Engle notes that a physician classifies a patient according to a diagnosis which necessarily subjugates or overlooks certain properties or aspects of the subject, and accents a single aspect around which the subject is classified (2). Although he does not state this as such, we may assume that this single aspect is illness or some manifestation of illness. In psychiatry, if there is one manifestation of illness, this one aspect is behavioral disturbance, using a broad definition for the term, "behavior." In contrast to the physical expression of illness; e.g., deviations in blood pressure, there is considerably more subtlety in distinguishing normality from abnormality in the spectrum of behavior manifested. Furthermore, the disturbed behaviors do not lend themselves easily to precise categorization. In our present state of psychiatric knowledge, therefore, it may not be possible to accent a single aspect or a similar parameter around which a subject can be classified. In fact, if one scans our present nomenclature, it is apparent that a number of parameters have been used, varying from one diagnostic category to another. Given signs and symptoms which lend themselves so poorly to measurement and grouping, and given a classification scheme with a potpourri of basic parameters, it is no wonder that we continue to be plagued by a low degree of reliability in psychiatric diagnosis.

It should also be noted, as observed in an editorial by Benjamin Pasamanick (4) that the unreproducibility of diagnosis may result more often from neglect rather than inherent unreliability in the diagnostic system. Although such neglect may be partially corrected by providing a more appropriate system of nomenclature, the proper use of any nomenclature will remain a matter of education and of cooperation by those using the system.

Through the years various psychiatrists have observed the deficiencies in existing classification systems, and have written to this effect and argued for revision. In tracing the development of statistics of mental disease in the United States, Horatio Pollock noted that in 1913 a committee of the American Medico-Psychological Association, in attempting to develop a uniform system of statistical reporting, urged the establishment of a standard nomenclature of mental diseases. According to Pollock, the situation among the States by 1917, in terms of the use of some system of classification, was chaotic (5). Stengel, writing in 1959, remarked that a " * * * serious obstacle to progress in psychiatry is difficulty in communication * * * The lack of a common classification of mental disorders has defeated attempts at comparing psychiatric observation and the results of treatment undertaken in various countries and even in various centers of the same country. Possibly, if greater attention had been paid to these difficulties, there might be a greater measure of agreement about the value of specific treatment than exists today" (8).

In the foreword to the "Diagnostic and Statistical Manual of Mental Diseases" of 1952, George Raines attested to the inappropriateness of the system then used, particularly for outpatient practice. "At the beginning of World War II, American psychiatry, civilian and military, was utilizing a system of naming developed primarily for the needs and caseloads of public mental hospitals. The origin of

this system was in itself predictive of the difficulties which would soon be encountered." By 1948 the situation in psychiatric nomenclature had deteriorated almost to the same point of confusion which existed throughout medical nomenclature in the 1920's. At least three nomenclatures (Standard, Armed Forces, and Veterans Administration) were in general use, and none of them fell accurately into line with the International Statistical Classification. In attempting to determine the sentiments of psychiatrists using the then (1948) current standard, it was learned that the high percentage of psychiatrists contacted felt that changes were urgently needed. "The need for change seemed to be felt more strongly by those in clinics and private practice than by those in mental hospitals or institutional work. However, a considerable proportion of mental hospital personnel urged change; this was especially true where outpatient clinics had been established in connection with the hospitals" (6). Thus, the problems in outpatient practice, as evidenced during World War II, led to our last major revision of the "Standard APA Nomenclature." All the handicaps to the utility of classification systems for the purpose of describing illness in the individual are only amplified in the outpatient practice of psychiatry where there exists, in addition to these handicaps, a variety of new problems unique to outpatient psychiatry. It is to these latter problems and the use of our classification scheme in outpatient practice that this paper will address itself.

Before addressing ourselves to the unique problems which outpatient practice of psychiatry presents to classification systems, let us first clarify what is to be subsumed under the term outpatient psychiatry for this presentation.

Outpatient practice encompasses an ever increasing range of psychiatric services. The report on comprehensive psychiatric programs published by the Joint Informa-

tion Service³ lists the following services in addition to the traditional clinic services or private practice: emergency service, aftercare clinics, consultative services for community agencies, formal community education programs, foster homes and nursing home care, ex-patient clubs, and recreational and vocational therapy programs. We might also add halfway houses, home visiting programs, and mental health teams to this roster. The gamut of patient groups seen in outpatient practice can be further extended by the addition of a variety of specialized clinics serving the courts, industries, alcoholics, geriatric patients, and adolescents. The broad spectrum of services, population groupings, and human behavior which fall within the ken of the psychiatrist in outpatient practice is emphasized by simply glancing at the above list. With the growth of community psychiatry and the attempts to provide more appropriate, effective psychiatric care to larger segments of the population, outpatient practice will become even more diffuse.

Broadly speaking, a system of classification serves both communicative and heuristic values. In classifying persons seen in the outpatient practice of psychiatry we might wish to communicate via diagnostic terminology a description of the person's illness. The current APA diagnostic system, however, is useful only in terms of gross categories, and the description generally communicated may or may not be appropriate to the individual at any particular phase of his illness. Secondly, the system outlines something of the cause of the illness, and thirdly, a prediction of the future course of the disease. The system is most limited in communicating the severity of an individual's current disability from his illness and a judgment about the response to any particular kind of therapeutic intervention.

³ The Joint Information Service of the American Psychiatric Association and the National Association for Mental Health: "Comprehensive Psychiatric Programs: A Survey of 234 Facilities," Washington, D.C., 1964.

In this paper, we will disregard the use of the classification system for research purposes (which are obviously not inseparable from any clinical usage of a classification system), and we will give our attention entirely to the value of a classification scheme in the decision-making processes of outpatient psychiatric practice.

What decisions, then, are required of the practitioner in outpatient psychiatric practice, and where in the course of this decision-making does diagnostic judgment play a role? Obviously, a diagnostic judgment is only one of many factors which influence the ultimate decisions that a practitioner must make about any particular patient. Equally obviously, this diagnostic judgment cannot be completely extricated from the complex of factors that lead to a particular decision. It is therefore necessary to remember that we are abstracting the one factor of diagnostic judgment from this context, and are subject to the artificiality characteristic of such a process.

Basically, there are four kinds of decisions which a practitioner must make: (a) evaluative; (b) therapeutic-interventional; (c) dispositional; and (d) consultative. The primacy of each of these decisions and the various factors influencing them will be determined in great measure by the setting in which the practitioner operates. In the usual setting of clinic or private practice, he is concerned largely with whether he should treat or not treat this person in that particular setting. If he decides that the person is suitable for treatment within the clinic or private office, he must then determine the kind of treatment which will be provided; i.e., the proper balance between somatic and psychotherapeutic, and which type of treatment under either of these two headings will be required. He will also be concerned with who is to be treated—will it be an individual or will it be a particular unit of people; for example, a retired couple, family, etc. The diagnostic system does not account for the multiperson factor; i.e., the concept of an individual as

part of a group and, therefore, in this context it is of no value in deciding whom to treat.

If the person is not suitable for treatment in his clinic or office setting, the practitioner's decision will be one of disposition. If care is required, he may first be concerned with whether this individual requires institutional care or whether he can remain in the community. If he can be maintained in the community, the practitioner must decide if the person needs a supervised setting such as a halfway home or perhaps some other kind of psychiatric or nonpsychiatric outpatient care.

Again, at only a very gross level is the diagnostic system indicative of the required degree of supervision. The practitioner is inclined to depend on the psychotic versus nonpsychotic dichotomy in reaching this type of decision. Yet, these categories do not suggest an appropriate disposition, except in a gross way; that is, hospitalization would probably be more readily recommended for a schizophrenic than for a non-schizophrenic.

In some settings, such as an emergency department or intake unit, dispositional judgments will absorb the greater focus, since the diagnostician will most often be seeing people early in the course of their illness. In other settings, such as aftercare clinics and transitional services, he will observe persons in the convalescent stages and will of necessity be more concerned with judgments related to therapeutic intervention and much less with any decisions regarding disposition. In these latter settings, he may be less concerned with assigning a particular category to the patient and more with the state of disability which this patient manifests. In both of these settings, he is involved with a particular phase of the person's illness, an aspect to be discussed below.

In the current practice of outpatient psychiatry, there are a number of other issues which must be considered if any system of classification is to be established that will

be adequate for the future. These issues will acquire even more significance if community or public health psychiatry assumes a greater role in determining the type of outpatient practice that will develop in future years.

Firstly, in psychiatry we are faced largely with chronic illness (often characterized by episodes of acute decompensation superimposed upon the more continuous symptomatic picture) and, thus, view fluctuations in symptomatology as a person progresses from the latent and early phases of his illness to the later more disabling stages. The observation of these phases of illness, or what is indeed the entire spectrum of illness, is more idiosyncratic to the outpatient practice of psychiatry than to any other area of psychiatry. This is particularly evident in the two current trends which are largely the province of outpatient psychiatry, one focusing on prevention and early case detection, and the other on rehabilitation and aftercare. The trend toward early diagnosis and prevention has meant that the patient is being seen during the phase of illness usually not observed by the diagnostician. For example, he may not be ill but potentially so, or only the early stages of the disease may be apparent. The practitioner must determine if the behavior presented is symptomatic of the prodromal phase of a more chronic reaction, or if it is a transitory disturbance unrelated to the progress of a particular disease. There is at present no means of classifying or diagnosing the illness in these early stages, except by means of terminology limited to defining symptoms extant in more advanced stages of the disease.

The consideration of this problem also throws into play the question of a gray area between health and illness or where one ends and the other begins. Because there is no category for normal behavior, the diagnostician is challenged to describe the condition as illness or as a variation of normal behavior. This question has arisen notably in child psychiatry and has added

to the argument for revision of the APA classification system (3).

Outpatient psychiatry focuses much attention on rehabilitation and aftercare. Yet there is no standard diagnostic terminology to define the patient's level of convalescence; our classification scheme does not take into account the social adjustment factors operative in recovery or the degree of social disability. A person treated for psychotic depression, for example, who is transferred to some kind of outpatient care, cannot then be diagnosed except in a language which defines the earlier and more acute stage of his illness.

A second problem which outpatient psychiatry presents to the usefulness of existing classification schemes is evident in the fact that it affords contact with population groups with whom we usually have little contact, and for whom the present classification systems are largely inappropriate. For example, among persons seen in a court clinic it is frequently difficult to determine whether the disorders which they present are personal or cultural in origin, and further, to classify one or both of these factors within the present systems. Home visits frequently make apparent dynamic differences between social and psychiatric problems; yet again there remains the impossibility of classifying them as working parts of the diagnosis.

Thirdly, the more traditional approach of "one person, one illness" is negated by the experience of outpatient psychiatry. In this setting, the patient is viewed within the larger focus of the multiperson dynamic; that is, the patient must be considered within the larger framework of his environment and the groups to which he belongs, be they family, parents, peers, etc. The problems which he presents may be more a product of this framework than of his own character structure. To attempt to accommodate present diagnostic terms to this consideration is a monumental task.

Given the range and heterogeneity of service provided in outpatient practice, and

given the variety of decisions that a practitioner is called upon to make in any specific outpatient setting, how useful is our current classification system? Is it of any use in arriving at a decision and does our present nomenclature serve any purpose in the communication of this decision?

At the University of Rochester Department of Psychiatry we have the unique opportunity to study psychiatric nomenclature through a psychiatric case register. We are thereby able to study the nomenclature at its operational level as it is utilized in a broad gamut of psychiatric services serving a wide range of psychiatric patients. This case register was initiated in January 1960 for the psychiatric services provided to the residents of Monroe County, N.Y. A comprehensive description of the goals and operation of the register has been presented elsewhere and only a brief account of the method will be provided here. Since January 1960, all but an estimated 3 percent of the county residents seen in the five psychiatric units, nine outpatient services, and by the 55 psychiatrists in private practice have been reported to the register. For each individual, a report is received for every episode of service, whether diagnostic or therapeutic.

The services that report to the register include almost a complete range of psychiatric facilities: a State hospital and its outpatient services, the inpatient, clinic, and emergency department services of a university hospital, an acute observation unit, a VA hospital and clinic, a children's residential treatment center and clinic, two child guidance clinics, and a court clinic. Psychiatric residents working in different services of the university hospital, the alcoholism and court clinics, and staff in all of the psychiatric services report to the register. Finally, 55 psychiatrists who spend varying proportions of their time in private practice report to the register; their training, range of experience, and nature of practice vary considerably.

Data are reported to the central file upon admission or initiation of service, and a search is made for name linkage. One or more diagnostic impressions may be given with the most acute disorder reported first. The standard APA nomenclature is used. All of the data are coded, punched on IBM cards, and later transferred to magnetic tapes for analysis by computer programs.

In order to view the use of our current nomenclature in outpatient practice, a tabulation was made of all the diagnostic terms used for all contacts with outpatient services during the years 1960 through 1963. A similar tabulation was repeated, this time excluding the State hospital and VA clinics. In these latter two services, patients are seen almost solely after hospitalization. As might be expected, a higher proportion of diagnostic terms contained in the category "psychosis" was found when State and VA hospital clinics were included. Otherwise, there was little change in the proportions, and the pattern was essentially the same for both tabulations. The proportion of the terms used within each major diagnostic category is shown in table I (pp. 41, 42).

In 74 percent of the contacts, only one diagnostic term was used, and in 3.5 percent of the contacts, three or more diagnoses were given. Nine diagnostic terms comprised 53 percent of all those assigned: psychotic depression, 2.7 percent; paranoid schizophrenia, 8.2 percent; chronic undifferentiated schizophrenia, 4.8 percent; anxiety reaction, 4.3 percent; neurotic depression, 6.5 percent; passive aggressive personality disorder, 9.8 percent; hysterical personality, 6.5 percent; adolescent adjustment reaction, 3.1 percent; and addiction alcoholism, 7.4 percent. Although the frequency with which some of these terms have been used is undoubtedly related to the frequency with which the disorder actually occurs in the outpatient population, it is obvious that some of the terms are used like wastebasket diagnoses.

TABLE I.—*Tabulation of Psychiatric Diagnoses for All Psychiatric Clinic and Private Office Contacts, Monroe County, January 1960 Through December 1963*

Diagnoses	Number	Percent
All categories	16,015	100.0
Acute brain disorders	183	1.14
Drug or poison	48	.30
Alcohol	29	.18
Acute hallucinosis	26	.16
Delirium tremens	38	.24
All other acute brain disorders	42	.26
Chronic brain disorders	607	3.79
Due to prenatal influence	1	.01
Due to syphilis	22	.14
Due to infection (nonsyphilitic)	10	.06
Drug intoxication	10	.06
Alcohol intoxication	80	.50
Due to trauma	21	.13
Cerebral arteriosclerosis	247	1.54
Senile brain disease	45	.28
Convulsive disorder	50	.31
Other conditions	39	.24
Unknown cause	82	.51
Mental deficiency	246	1.54
Psychotic disorders	4,158	25.95
Involutional psychotic reaction	307	1.92
Psychotic depressive reaction	427	2.66
Other affective reactions	303	1.89
Schizophrenic reaction and chronic undifferentiated schizophrenia	762	4.75
Schizophrenic reaction:		
Paranoid type	1,321	8.24
Schizo-affective type	182	1.14
Catatonic type	400	2.50
Acute undifferentiated type	59	.37
Hebephrenic type	93	.58
Simple type	38	.24
Residual and childhood type	35	.22
Paranoid reactions	164	1.02
Other psychosis	67	.42
Psychophysiologic disorders	65	.41
Psychoneurotic disorders	2,591	16.17
Anxiety	682	4.26
Dissociative	20	.12
Conversion reaction	239	1.49
Phobic reaction	158	.99
Obsessive reaction	122	.76
Depressive reaction	1,039	6.48
Others	331	2.07
Personality disorders	5,689	35.50
Inadequate personality	157	.98
Schizoid personality	140	.87
Cyclothymic personality	13	.08
Paranoid personality	70	.44
Emotionally unstable personality	103	.64
Passive aggressive personality	1,569	9.79
Hysterical personality	1,038	6.48
Compulsive personality	112	.70
Sociopathic personality	166	1.04
Sexual deviation	199	1.24
Addiction, alcohol	1,192	7.44

TABLE I.—*Tabulation of Psychiatric Diagnoses for All Psychiatric Clinic and Private Office Contacts, Monroe County, January 1960 Through December 1963—Continued*

Diagnoses	Number	Percent
Personality disorders—continued.		
Addiction, drug	26	.16
Special symptom reactions	91	.57
Others (not specified)	813	5.07
Transient situational disorders	1,086	6.78
Adult situational reaction	176	1.10
Childhood adjustment reaction	391	2.44
Adolescent adjustment reaction	502	3.14
Others	16	.10
Diagnosis deferred	1,113	6.95
Without mental disorder	162	1.01
Other diagnoses (anorexia nervosa, suicide attempt)	115	.76

A more detailed examination of the individual diagnostic terms and the frequency with which they are used suggests several modifications to the current nomenclature. The diagnoses subsumed under the categories “acute brain syndrome” and “chronic brain syndrome” are used rarely, and this is thought to be a function of patterns of service in that most of these people are seen by the inpatient services. This thesis is supported by a similar tabulation of all the inpatient admissions. Of the 21,288 diagnoses assigned to hospitalized patients from 1960 through 1963, 2,818 (13.2 percent) were in the category, “acute brain disorder,” and 3,321 (15.1 percent) fell into the category “chronic brain disorder.”

Under the acute brain syndrome category, one wonders about the detailed categorical divisions provided. Except for the diagnosis of acute alcoholism and acute drug intoxication, the other terms were seldom used. This was true for both the inpatient and outpatient services. Perhaps greater specificity would promote better utilization of the terms in this category. In addition, we may need further education of the psychiatric profession if we are to make use of the more specific terminology. Psychiatrists may well be missing or ignoring the milder cases of acute brain syndrome and focusing most of their attention upon the psychological conflicts. With

the ever-increasing use of medication it is surprising to see the diagnosis, acute brain syndrome due to drug intoxication used for only 48 patients out of a population of 16,015. It might be helpful to revert to more gross categorization and use acute brain syndrome as a qualifying diagnosis with the other conditions; i.e., intracranial neoplasm, metabolic disorder, convulsive disorders, etc. In the chronic brain syndrome category, one might question some of the specificity attempted. It would appear that one term for a chronic brain syndrome associated with central nervous system syphilis would suffice, and that the degree of differentiation sought in the present nomenclature is outmoded. Similarly, the degree of differentiation sought for chronic brain disorder associated with trauma may be more than is warranted. One or two diagnostic terms might be sufficient. During the 4-year period, 21 cases of chronic brain syndrome associated with any type of brain trauma were reported from the outpatient services (34 from inpatient services).

The use of the diagnosis “mental retardation” could not be evaluated by means of the psychiatric case register since these cases are only reported to the register when associated with some other psychiatric condition. No facilities for the care of the mentally retarded report to the Monroe County case register.

The nomenclature for the category "psychotic disorders of psychogenic origin" was used irregularly. Under affective reactions, only the diagnosis, psychotic depressive reaction, was used with any frequency. Manic-depressive diagnoses were used infrequently, and then, only by psychiatrists practicing at our State hospital.

For the schizophrenic reactions, four terms were used for almost all of the cases: "chronic undifferentiated schizophrenia" or simply "schizophrenic reaction" as a nonspecific term; "paranoid schizophrenia"; "schizoaffective reaction"; and "catatonic schizophrenia." The diagnoses, simple and hebephrenic schizophrenic reactions, which are seldom used and then only by our State hospital psychiatrists, are generally assigned to patients who were admitted to State hospital several years ago, and who were assigned these two diagnoses at that time. Surprisingly enough, the term "schizophrenic reaction, residual type" is rarely used, though it might be thought that there is a definite need for a diagnosis of this type. Under the paranoid reactions, little attempt was made to distinguish the type of reaction. The terms paranoia or paranoid state are rarely used and it would seem that one diagnostic term would be sufficient for this type of reaction.

The category "psychophysiological disorder" appears to be one that has not become popular with the practicing psychiatrists. One of these diagnoses was used in only 64 cases and in half of the cases the diagnosis, psychophysiological disorder, gastrointestinal type, was used. One could question whether this should be retained as a major category. The generic term itself might be sufficient and the specific type of disorder might be recorded as a qualifying phrase.

There is little to be said regarding usage of the diagnostic terms under the category "psychoneurotic disorder." Most of the terms were used frequently and the one term used infrequently, "dissociative reaction," appears to be due more to the

rarity of the disorder than to any difficulty with the diagnostic term itself.

As noted above, two diagnostic terms, "passive aggressive personality" and "hysterical personality," were used like wastebasket terms for the category "personality disorder." These two diagnoses were utilized in 65 percent of the cases in which a diagnosis of personality pattern or personality trait disturbance was given. Passive aggressive personality was used five times more frequently for men than women, and the reverse was true for the term "hysterical personality." Under personality pattern disturbance, the terms "cyclothymic personality" and "paranoid personality" were used infrequently. The description of cyclothymic personality as provided in the "Diagnostic and Statistical Manual" may not provide a sharp enough differentiation from the "normal" personalities. In contrast, the description provided for the paranoid personality may approach too closely the definition given for paranoid reaction or paranoid schizophrenia, and might thereby account for the infrequency with which this term is used. The term "sociopathic personality" appeared to be used most often in covering the three diagnostic terms provided in the nomenclature. There seemed to be little attempt to distinguish between antisocial reaction, dy-social reaction, or simply sociopathic personality. Nonspecific terms such as "personality trait disturbance," "other," "mixed personality trait disturbance," and simply "personality trait disturbance with no qualification," were used quite frequently, further signifying the inadequacy of this diagnostic category to describe anything beyond the gross levels of differentiation.

There is little that needs to be said about the category "transient situational personality disorders." The diagnosis, adjustment reaction of late life, was rarely used, perhaps due to the difficulty in distinguishing this from other diagnoses used for this period of life.

We have discussed at some length the kinds of decisions that must be reached in outpatient psychiatric practice and the role that diagnostic judgment might play in such decisions. As we have already stated, factors that may influence any administrative decision in outpatient practice or any decision concerning therapeutic intervention are numerous and complex. It may be presumptuous to expect that we can isolate the part which diagnostic judgment plays in this complexity. Therefore, we thought that it would be useful to obtain at least some glimpse of the association between diagnostic and dispositional judgments in outpatient practice. The psychiatric case register has given us some clues concerning this association.

In order to view this relationship we selected all those patients contacting the outpatient services of Monroe County during the years 1961 and 1962. We chose only those persons who had had no psychiatric contact for at least 1 year or more prior to the outpatient contact in 1961 or 1962. We further decided to examine only contacts with what might be called the community services, including private office practice, and to exclude those persons contacting only the State hospital or VA clinics. The latter groups of persons were those who had been discharged from the hospital and were entering the clinic for after-care, thus limiting the number of decisions that would be open regarding their therapy or further disposition. Finally, we limited the group to adult patients ages 15 and over, and split the entire cohort into two subcohorts: those who had a history of previous psychiatric hospital care and those who had no history of psychiatric hospitalization.

During this 2-year period, 5,087 patients age 15 and over contacted one of the community outpatient psychiatric services. At the time of contact with the outpatient service, the professional staff made decisions concerning possible therapeutic intervention or referral to some other source of

help. Essentially four choices concerning psychiatric intervention were available to the practitioner, and accessible to study through the psychiatric case register: (a) No psychiatric intervention is necessary; (b) outpatient psychiatric treatment would be helpful either at that same clinic or by that same practitioner; (c) outpatient psychiatric care would be helpful, but at another clinic or by another practitioner; or (d) inpatient psychiatric care is necessary. The decision to treat the individual could be subdivided further into a decision to provide only short-term therapy (arbitrarily designated as less than 3 months), or a decision to provide long-term therapy (longer than 3 months).

Through the psychiatric register we were able to study the psychiatric service, if any, that the entire group of patients received beyond their initial outpatient contact. We have arbitrarily categorized these patients into five groups according to the pattern of psychiatric service they received. For the sake of initial simplicity, at least, we have assumed that this was the type of service recommended by the practitioner. Obviously, this overlooks such factors as the patient's acceptance of treatment, his financial status, pressures by family members and other forces varying the kinds of care such as the availability of beds, the availability of a therapist in an outpatient facility, and the many other factors that play a role in the type of service which any patient receives. The five groups are as follows:

1. Those who had only diagnostic service and who had no subsequent psychiatric contact for at least 1 year or longer.
2. Those who received treatment at the time of their initial outpatient contact and who remained or were kept in treatment for a period of less than 3 months.
3. Those who received treatment at the time of their initial outpatient contact and remained in treatment for 3 months or longer.

4. Those who received only diagnostic service at the time of their initial outpatient contact and who were seen in private office practice or in one of the clinics within a period of 15 days or less.

5. Those who received diagnostic service at the time of their initial outpatient contact and were admitted to an inpatient psychiatric service within a period of 7 days from the time of their original outpatient contact.

If a diagnostic judgment enters into the decision regarding the type of care that a patient requires, then we would expect to see a considerable variation in the pattern of service received according to the various diagnostic categories. Tables II and III (pp. 46 and 48) show the pattern of psychiatric service received by diagnostic category for (a) those persons with no history of prior psychiatric hospital care, and (b) those with a history of prior psychiatric hospital care. Obviously, we cannot say from this material whether the clinical judgment regarding therapeutic intervention and disposition influenced the diagnostic judgments. At this point we can examine only the associations between the pattern of service and diagnostic judgment and speculate about these associations.

First, we might expect that those persons with a diagnosis of psychotic reaction, with chronic brain syndrome, or with some acute reaction, would be hospitalized more often than those not diagnosed as such. In looking at table II, we find that this thesis is supported; approximately one-fifth to one-third of those persons with psychosis or with chronic brain syndrome were hospitalized initially. Persons diagnosed as neurotically depressed were hospitalized at a greater rate than the other categories but less than the psychotics. The 6 percent hospitalized under the category of "other diagnoses"⁴ resulted largely from the

individuals who received a diagnosis of acute brain syndrome.

Similarly, we might expect that those receiving a diagnosis of personality disorder, or any other term indicating a benign condition, would be less apt to receive psychiatric care than those falling within other categories. Again, this proves to be the case in that approximately 40 percent of the persons with personality disorder or a transient situational personality disorder received only the initial diagnostic service and had no further psychiatric care. Three-fourths of those in the group "other diagnoses" received only the initial diagnostic evaluation. These were primarily persons who were said to be without mental disorder or for whom the diagnosis was deferred.

An exception to this pattern of greater rates of treatment for psychotics and greater proportions receiving only diagnostic service in the nonpsychotic population was noted in the chronic brain syndrome group. Here, 55 percent received only an initial diagnostic contact. Although approximately one-fourth of all those with chronic brain syndrome were hospitalized, the diagnostic judgment for the majority of persons falling within this category seemed to have less relationship to the ultimate type of service received than is true for other diagnostic categories. There appears to be more of a continuum of illness and disability in the schizophrenic disorders than in the chronic brain syndrome group. In the latter group there appears to be a dichotomy rather than a gradation between those who present severe problems and those who do not. This contrast between the two groups may also lie in the continuum of services available for schizophrenics in contrast to the psychiatric services provided for those with chronic brain syndrome. The outpatient services, in particular, seem more appropriate for the schizophrenics.

⁴ Hereinafter, the term "other diagnoses" is used to describe a category which we have arbitrarily designated to include suicide attempt, diagnosis deferred, and persons listed as without mental disorder.

TABLE II.—All Individuals, Age 15 Years and Over, Admitted to Outpatient Psychiatric Service (Clinic and Private Office Practice, Excluding State Hospital Clinic and Veterans Administration Clinic) in 1961 and 1962

No history of psychiatric inpatient care and no outpatient psychiatric contact for at least 1 year prior to 1961 or 1962 outpatient contact

[Percent distribution by diagnoses and disposition after the 1st contact]

Dispositional subgroups	Total		Schizo-phrenic	Affective psychoses	Chronic brain syndrome	Neurotic depressive	Other neuroses	Alcoholism	Situational reaction	Personality disorders	Other diagnoses ¹
	Number	Percent									
I. No treatment, no immediate disposition	1,470	37.1	20.0	10.1	55.2	28.1	28.4	25.1	43.8	42.7	72.5
II. Treatment, less than 3 months.....	1,044	26.4	27.1	30.9	11.0	32.4	25.8	35.7	32.4	25.5	12.8
III. Treatment, 3 months or more.....	1,020	25.8	26.8	20.8	9.7	28.6	39.1	33.3	20.6	25.5	6.1
IV. No treatment, transfer to other outpatient service.....	117	3.0	5.1	1.7	.7	2.3	3.3	2.4	2.6	3.4	2.7
V. No treatment, transfer to inpatient service....	306	7.7	21.0	36.5	23.4	8.6	3.4	3.5	.6	2.9	5.9
Total.....	3,957	100.0	100.0 N = 295	100.0 N = 178	100.0 N = 154	100.0 N = 676	100.0 N = 701	100.0 N = 171	100.0 N = 349	100.0 N = 1,058	100.0 N = 375

¹ Other diagnoses include suicide attempt, diagnosis deferred, and without mental disorder.

In all of these categories except "chronic brain syndrome" and "other diagnoses," approximately 50 to 65 percent of the entire group received some kind of outpatient treatment on their initial contact. Those persons with a diagnosis of neurosis other than neurotic depression had a greater tendency to remain in treatment longer than 3 months. In contrast, for all diagnostic categories except "other neuroses" and "other diagnoses" there was a slightly greater tendency to remain in treatment less than 3 months. It would appear that except for the diagnosis, neurotic reaction, diagnostic judgment plays little role in the decision regarding duration of outpatient treatment. For the diagnosis, neurotic reaction, there is a greater tendency to remain under care longer than 3 months, whereas there is little tendency for those with a diagnosis, chronic brain syndrome, to be kept in treatment.

In table III the same comparison of diagnostic category by pattern of service received is shown for the adult outpatients who had a history of prior hospital care at the time of their outpatient contact. In comparing tables II and III we can see the influence that a history of prior hospital care, indicative of chronicity, may have had upon determining the pattern of service provided for a patient. Those with a history of prior hospital care are much less apt to have only an initial diagnostic contact, and a much greater chance of receiving some kind of psychiatric treatment. In all of the categories other than "affective psychosis," those with a history of prior hospital care had a greater rate of hospitalization than those without such a history. This difference is most noticeable for the categories that include the nonpsychotic disorders. Although the practitioner has assigned a diagnosis of nonpsychosis, when there is a history of prior inpatient care, his thoughts about the severity of illness and the need for further hospital care now become paramount in determining this decision. This history of prior hospital

care seems to play less of a role for those with a diagnosis of psychosis. With a diagnosis of affective psychotic reaction, it may be that a repeat episode is viewed with pessimism in terms of the possible outcome of therapeutic management and hospitalization; here, fewer persons are hospitalized when there is a history of previous inpatient care. Again, for the group with a history of prior hospital care, the diagnosis seems to play little role in determining outpatient treatment.

Approximately 65 to 70 percent of the persons in each diagnostic category received some kind of outpatient treatment. A somewhat greater proportion of those with a diagnosis of alcoholism received outpatient care, and this may be attributed to the special clinic for alcoholics within the complex of outpatient services in this area. Again, those within the category "other diagnoses," largely with diagnosis deferred, were much less apt to receive outpatient care.

As we have noted before, many factors play a role in determining the ultimate pattern of care that any individual receives. For example, if we find that a larger proportion with a diagnosis of psychosis are hospitalized than those without such a diagnosis, this apparent relationship may be due to a third factor related to both variables. Socioeconomic status, type of family pattern, or a number of other influences may affect both the rate of hospitalization and the frequency with which any group of patients are assigned the diagnosis of psychosis.

To view the influence of socioeconomic status, we have compared the distribution of service patterns according to diagnostic category and area of residence. Thus in table IV (p. 49) we have selected those individuals with no history of prior hospital care who were diagnosed as schizophrenic and those with a diagnosis of some type of personality disorder. For each diagnostic category, we compared the pattern of service received for persons residing

TABLE III.—All Individuals, Age 15 Years and Over, Admitted to Outpatient Psychiatric Service (Clinic and Private Office Practice, Excluding State Hospital Clinic and Veterans Administration Clinic) in 1961 and 1962

With history of previous psychiatric inpatient care and no psychiatric contact for at least one year prior to 1961 or 1962 outpatient contact

[Percent distribution by diagnoses and disposition after the 1st contact]

Dispositional subgroups	Total		Schizo-phrenic	Affective psychoses	Chronic brain syndrome	Neurotic depressive	Other neuroses	Alcoholism	Situational reaction	Other diagnoses ¹	Personality disorders
	Number	Percent									
I. No treatment, no immediate disposition	142	12.6	8.6	2.4	19.4	14.9	20.9	8.0	20.0	45.3	17.1
II. Treatment, less than 3 months.....	361	31.9	27.8	35.7	35.5	38.6	33.3	36.3	45.0	15.1	32.2
III. Treatment, 3 months or more.....	389	34.4	37.7	35.1	13.0	28.3	33.3	41.6	20.0	17.0	38.2
IV. No treatment, transfer to other outpatient service.....	30	2.7	2.3	.6	0.0	1.7	3.9	1.7	10.0	9.4	4.6
V. No treatment, transfer to inpatient service...	208	18.4	23.6	26.2	32.1	16.5	8.6	12.4	5.0	13.2	7.9
Total.....	1,130	100.0	100.0 N = 385	100.0 N = 168	100.0 N = 31	100.0 N = 127	100.0 N = 81	100.0 N = 113	100.0 N = 20	100.0 N = 53	100.0 N = 152

¹ Other diagnoses include suicide attempt, diagnosis deferred, and without mental disorder.

TABLE IV.—All Individuals Receiving a Diagnosis of Schizophrenia or Personality Disorder, Age 15 years and Over, Admitted to Outpatient Psychiatric Service (Clinic and Private Office Practice, Excluding State Hospital Clinic and Veterans Administration Clinic) in 1961 and 1962

No history of previous psychiatric inpatient care and no psychiatric contact for at least 1 year prior to 1961 or 1962 outpatient contact

[Percent distribution by area of residency and disposition after the 1st contact]

Disposition subgroups	Schizophrenia				Personality disorder			
	Inner city		Suburban		Inner city		Suburban	
	Number	Percent	Number	Percent	Number	Percent	Number	Percent
I. No treatment, no immediate disposition.....	8	36.4	19	24.7	47	59.5	124	40.6
II. Treatment, less than 3 months...	7	31.8	25	32.4	19	24.1	74	24.3
III. Treatment, 3 months or more....	4	18.2	22	28.6	11	13.9	98	32.1
IV. No treatment, transfer to in-patient service.....	3	13.6	11	14.3	2	2.5	9	3.0
Total.....	22	100.0	77	100.0	79	100.0	305	100.0

in the inner city area with those whose residency is within the suburban area. This is a crude index of socioeconomic status, and all we can really say is that we are comparing patterns of service to persons who reside within a very low income area with those in a middle to high income area of the county.

We may note that there was a greater tendency for schizophrenics to be hospitalized than for those with a diagnosis of personality disorder, and for the latter group to receive the initial diagnostic service with no subsequent care. However, within each diagnostic category, there is a greater tendency for persons from the inner city than for those from the suburban area to receive only the initial diagnostic service. These differences are more striking for the personality disorder groups than for the schizophrenics. We may also note that within each of these two diagnostic categories, the percentage hospitalized is approximately the same from the two different areas. The individual's place of residence (socioeconomic status) seems to play a role in whether or not a person will receive any outpatient care. Although financial status may be a determining factor in this pattern, it cannot be the sole explanation since most of the persons from the inner city were seen in clinic

contacts and therefore, presumably, could have been maintained in treatment.

However, given a certain level of illness, as partly indicated by the diagnostic judgment; e.g., schizophrenic versus personality disorder, factors such as the person's area of residence became less important in determining the service received. Given certain levels of illness, as indicated by the diagnosis assigned or even within a particular diagnostic category, a patient was more apt to be hospitalized or to receive at least some kind of care. The pattern remains the same both for those with a history of prior hospital care and those without such a history, as shown in table V (p. 50).

Summary

What can be said, therefore, about our current APA nomenclature as it is utilized in outpatient psychiatric practice and what changes might we recommend? It is evident that the present nomenclature may figure in the decision-making processes of outpatient practice. It may permit some gross differentiation of patients: those who require hospital care, those who need some other type of psychiatric care, and those who need no psychiatric care. Some of the diagnostic terms, i.e., "schizophrenic re-

TABLE V.—All Individuals Receiving a Diagnosis of Schizophrenia or Personality Disorder, Age 15 Years and Over, Admitted to Outpatient Psychiatric Service (Clinic and Private Office Practice, Excluding State Hospital Clinic and Veterans Administration Clinic)

With history of previous psychiatric inpatient care and no psychiatric contact for at least 1 year prior to 1961 or 1962 outpatient contact

[Percent distribution by area of residency and disposition after the 1st contact]

Disposition subgroups	Schizophrenia				Personality disorder			
	Inner city		Suburban		Inner city		Suburban	
	Number	Percent	Number	Percent	Number	Percent	Number	Percent
I. No treatment, no immediate disposition.....	6	12.8	6	5.8	4	21.1	7	17.1
II. Treatment, less than 3 months....	13	27.7	29	27.9	7	36.8	17	41.5
III. Treatment, 3 months or more....	15	31.8	47	45.1	6	31.6	13	31.7
IV. No treatment, transfer to in-patient service.....	13	27.7	22	21.2	2	10.5	4	9.8
Total.....	47	100.0	104	100.0	19	100.0	41	100.0

action,” “affective reaction,” and “chronic brain syndrome,” provide some impression of the degree of disability, though only for a limited segment of the population within each of these categories. In this sense, the diagnostic judgment appears to be associated with the decisions regarding the extent of therapeutic intervention. Obviously, we cannot be certain at this point whether the clinical judgment determines the diagnostic judgment or vice versa. Probably each influences the other. At levels of illness in which therapeutic intervention may be more elective than necessary, diagnostic judgment seems to play a much lesser role. This is particularly evident in the determination of the nature, duration, and intensity of outpatient care. Some diagnoses, particularly in the psychoneurotic category, do seem to influence both the decision to administer outpatient care and its duration.

Much of the nomenclature is not used or is used quite infrequently in outpatient psychiatric practice. In some parts of the standard nomenclature, particularly in the chronic brain disorders, and perhaps in the acute brain disorders, it would seem that there is a greater degree of specificity than is required or appropriate. A few diagnostic terms are used to represent entire categories, thus blurring the distinction

between any subdivisions within that particular category. This is notably demonstrated in personality disorders.

Many diagnostic terms within our standard nomenclature may well represent or symbolize particular syndromes and thus indicate a specific picture of illness. There are few terms that signify a particular phase of any disorder, and most of the diagnostic labels are representative of chronic disorders with wide fluctuations in symptomatology and levels of disability. Some of the terminology is based upon known or supposed causation. As we have noted, the dimension along which any group of disorders has been classified varies from one category to another with an etiologic or descriptive axis being the most common. A few of the categories and a variety of the diagnostic terms are indicative of prognosis, severity of disability, response to therapy, or the need for supervision. These factors, however, while they play a role in decisionmaking processes of outpatient practice, have either been ignored in our standard nomenclature or are implied only by general diagnostic terminology.

What changes might we recommend that would provide a classification scheme more suitable for outpatient psychiatric practice? For the sake of simplicity, the following proposals have been focused around out-

patient practice only, and thus have ignored other needs upon which a psychiatric classification should be based.

A multidimensional classification scheme would appear to be the ideal solution to our current dilemma; each patient would be classified along three or perhaps four axes. First, the patient would be classified according to the type of reaction or disorder which he manifests and, further, by the phase of illness. For example, a person might be classified as paranoid schizophrenic reaction, latent phase, or psychotic depressive reaction, convalescent stage. Secondly, the individual's level and course of disability might be indicated, the former by very gross categories such as "mild," "moderate," or "severe impairment." Again, the level of disability might be represented by more well-defined categories indicating whether the individual was able to work or could care for himself; whether he was potentially violent or potentially self-destructive. Concomitant with the gradation of disability, the course of the patient's disability might be indicated by such terms as "recent," "chronic," or "fluctuating." The categorization of the course of illness and disability level with a sufficiently concise terminology poses a difficult task. We do not believe this is an insuperable challenge, however, and even a relatively primitive classification of these variables may prove to be valuable.

Finally, an individual would be classified according to his basic pattern of adjustment or his basic character structure by terms such as "withdrawn," "depressive," "paranoid," "hysterical," and "antisocial type." We have excluded certain parameters such as premorbid adjustment, response to therapy, and others, as listed in the Department of Defense "Disease and Injury Code" used by the armed services. These dimensions are not easily classifiable and make for an unwieldy scheme.

Under such a multidimensional system, an individual might be classified as follows:

(a) undifferentiated schizophrenic reaction, asymptomatic phase of fluctuating disorder; (b) moderate disability, chronic; (c) withdrawn type.

It is also recommended that we consider the development of a new category in which we might include interactional disorders. Here we would classify in terms of groups or units rather than individuals, and would consider the disorders as relevant to married couples, family units, and a number of other small groups. Disturbances in communication and in interpersonal relationships would be considered within the fabric of the social settings in which they exist. The category "transient situational personality disorder" is a partial though inadequate move in this direction. Admittedly, this will be a difficult classification category to construct, but one which we believe is necessary, particularly in outpatient practice. Our inability to accommodate an individual-based classification scheme to the concept of an individual who is seen as part of a group, together with the trend toward small group therapy, signifies the need for this type of classification.

Since the ideal system is one to be sought but may be long in coming, we offer some recommendations about the current nomenclature, recognizing that these changes will be of minor importance in terms of utility of this nomenclature in outpatient practice. It is recommended that much less specificity be sought under the chronic brain disorders. The acute brain disorders require considerable thought and revision. Greater specificity and more detailed terminology might be attempted. If not, then only the terms for drug and alcohol intoxication should be kept under acute brain disorders, and all other diagnoses should be grouped into the one term, acute brain disorder, other. In the chronic brain syndrome category, only one term for chronic brain syndrome due to syphilis, and one or two for chronic brain syndrome

due to trauma, should be retained. We would question the need for the term, simple schizophrenic reaction in the category, "psychotic disorders," and would recommend that only one term be kept for paranoid reaction. Finally, the personality disorder category needs to be restudied. We believe that a more realistic and practical nomenclature could be developed for this category in contrast to its counterpart in our current standard nomenclature.

REFERENCES

- (1) BABIGIAN, HAROUTUN M.; GARDNER, ELMER A.; MILES, HAROLD C.; and ROMANO, JOHN: "Diagnostic Consistency and Change in a Followup Study of 1,215 Patients"; *American Journal of Psychiatry*, 121: 895-901, March 1965.
- (2) ENGLE, RALPH, and DAVIS, B. J.: "Diagnosis: Present, Past and Future," part II; *Archives of Internal Medicine*, 112: 520-29, 1963.
- (3) LANGFORD, WILLIAM S.: "Reflections on Classification in Child Psychiatry, * * *"; *Diagnostic Classification in Child Psychiatry*; American Psychiatric Association Research Report, 1964, pp. 1-21.
- (4) PASAMANICK, BENJAMIN: "On the Neglect of Diagnosis"; *American Journal of Orthopsychiatry*, 33: 397-8, 1963.
- (5) POLLOCK, HORATIO M.: "Development of Statistics of Mental Disease in the United States During the Past Century"; *American Journal of Psychiatry*, 102: 6-7, 1945-46.
- (6) RAINES, GEORGE W.: Foreword, "Diagnostic and Statistical Manual of Mental Disorder," 1952, pp. vi-viii.
- (7) SIMPSON, GEORGE G.: "Principles of Animal Taxonomy," Columbia University Press, New York, 1961, p. 3.
- (8) STENGEL, E.: "Classification of Mental Disorders"; *Bulletin of the WHO*, 21: 601, 1960.

Supplementary Comments by Dr. Gardner as Presented at the Conference

The structure of any classification scheme should vary with the objectives for which it is established. Therefore, I want to emphasize that my remarks pertain primarily to a certain area of psychiatry, to a particular population group, and to the role of the

psychiatrist as clinician and as administrator, although obviously this is not divorced from research and the role of the investigator.

As Dr. Brill has noted, the classification scheme serves largely a community role. Ideally, the assignment of a patient to a particular diagnostic category would permit us to make some prediction about his future course and, perhaps, his response to particular therapeutic programs. The diagnostic terminology would then allow us to communicate this judgment as well as a gross picture of the patient's disorder to others. As noted by others in the conference, a classification scheme can serve only a limited number of objectives. However, the current psychiatric classification, at least as I see it, falls far short of the ideal and only irregularly provides the kind of communication I have just noted.

In addition to all of the problems that we face in hospital psychiatry and in investigative work, outpatient practice creates a number of somewhat unique obstacles to the design of a useful classification scheme. I would like to comment about a few of these.

First, comparable to the situation in epidemiological work, a much broader spectrum of illness is seen in outpatient psychiatry than in other areas, making the differentiation of disorders even more difficult. The emphasis on early case detection and the contact with a wider sample of the population at large acquaint the clinician with many new behavioral patterns.

Secondly, outpatient practice encompasses an ever-increasing variety of services. The clinician's orientation may change from one setting to another and his role may vary. At one time he acts as a therapist and the next moment as a diagnostician, a consultant, or even as a presumed expert on human behavior. He must communicate with other medical practitioners and very often with nonmedical personnel.

With increasing horizons he comes into contact with problems or disorders that are not clearly confined to one person but often represent disturbances in social interaction, sometimes a result of individual personality patterns and at other times as much attributable to the social setting in which the interaction occurs as to the intrapsychic problem. Therefore, the significant unit of classification may change. At one time, it may be the individual and at another time the group or interactional processes.

Finally, the outpatient psychiatrist is particularly concerned with the degrees of disability. Obviously, this is a concern of all areas of psychiatry, but it well may loom even greater in outpatient psychiatry since here the patient is not in a protected or supervised setting and the judgments about his impairment may be more complex and subtle. For example, he may be asked whether an individual psychiatric disorder makes him a suitable parent and whether or not he should have custody of his children.

I realize that it seems presumptuous to suggest any classification scheme at this time and any such proposal should be quite explicit. As I have given thought to the problems of classification in outpatient practice it repeatedly seemed to me that a multidimensional scheme would be more useful than our present system, at least for outpatient work, and probably for investigative work. Such a system might be comparable to the American Heart classification. I believe that it would be feasible for psychiatry. For example, a patient would be assigned one or more diagnoses to describe the reaction or disorders noted in the present classification but then with terminology to designate the phase of illness which he presents; secondly, the level and course of his disability would be classified; and finally, his basic personality or adjustment pattern and, perhaps his response to some specific therapy, should be categorized. Obviously, this radical a change should be presented in more detail than is feasible at this time. The possibility of categorizing

or scaling these parameters must also be shown and, of course, the practicality of any system must be considered.

However, given the complexity of human behavior, how much can we expect any single, all-inclusive categorization to accomplish? I also might add that in outpatient practice currently, most clinicians already use our classification scheme as a multidimensional system. Many categories are used only at represented points along a single dimension but without any standardization.

There is one final point I would like to make:

I cannot help but believe that some of the attempts to develop new typologies and the use of multivariate analyses stem from a failure of clinicians to analyze the clinical process. We fail to make explicit the bits of data that actually enter into our diagnostic judgment. Perhaps, even more importantly, we make no attempt to weight the different factors just as we do in our day-to-day practice.

A Plan for Refining the Nosology of Mental Illness

Paul T. Wilson, M.D.¹

Refining a nosology involves much more than just renaming illnesses from time to time. The nosological system of any healing profession is an integral part of its structure, tradition, philosophy, and practice. No single part can advance far without similar growth in the other parts. And so the act of improving a nosology is inseparable from the other acts that help the profession to grow.

¹ Principal Investigator, Information Processing Project, American Psychiatric Association, Washington, D.C.

This paper was prepared under Grant MH-09219-02, National Clearinghouse for Mental Health Information, NIMH.

The Function of a Nosology

The ultimate purpose of a nosology is to help treat people who have a particular kind of disorder and to help others avoid it. All such acts of treatment and prevention have one thing in common: They are chosen to achieve the best results for the particular situation in which they are used. But such situations occur in an infinite variety of forms, and there are only a finite number of responses available to deal with them. The problem is solved, of course, by observing many of these situations and from these observations creating a series of classifying abstractions called by such names as "diseases," "syndromes," and "pathological processes." The professional person has only to classify the situation that confronts him and select the most appropriate response to it.

In order to do this he must know two things: How to recognize each class of situation, and what response that class of situation calls for. People entering a profession gain this information through the double process of education and training. Education, as I use the term here, is primarily a verbal process by which students learn through lectures and reading how to classify disorders, how to treat them, and how to recognize them. Because education is a verbal process, however, even a well educated student cannot, when he first encounters them, recognize most disorders without first translating what he sees into words and then recognizing in these words the disorder the patient has. (This is a limitation familiar to everyone who remembers the embarrassing transition from the pre-clinical to the clinical phase of his medical school studies.) What is needed to convert a student into a clinician is the process of training, in which he learns what stimuli to associate with the words he has learned.

The major function of a nosology (and the technical vocabulary that surrounds it) is to label abstractions that classify situations associated with certain disorders in

order to deal with those situations more effectively. As a collection of words, a nosology constitutes the building units of education. Students learn to use these words in making statements that help them to understand, to predict, and—as a result—to control those situations. A nosology also provides the framework around which trainees learn to recognize in real situations the abstractions they learned as students.

Another function a nosology performs is to facilitate the study of the disorders it classifies by serving as a kind of stable semantic nidus around which new facts may be gathered. For example, only when two people have a word for a disorder and agree on what that word means, can they instruct each other by reporting new findings about that disorder. Without the shared meaning of that word they might tell each other about two different diseases.

Prerequisites for an Effective Nosology

The ultimate test of a nosology is the effectiveness of the people who use it. This effectiveness depends both on how well the nosology is prepared and on how well it is used. To meet both criteria, a nosology should be (1) scientifically accurate, (2) realistically useful, and (3) consistently applied.

(1) "Scientific accuracy" refers to the relationship between the way events actually occur together and the way the meanings of our words assume them to occur together. In a scientifically accurate vocabulary, there should be a name for every significant group of events that occur together; it should leave no significant groups unnamed; and none of its words should refer to groups of events that are not significant. I will say more later about "significance."

(It is important to separate here the accuracy of a vocabulary from the accuracy of statements using the vocabulary. For example, the term "anxious," as it is usually defined, can be considered accurate because the events comprising anxiety occur

together consistently enough to justify using the word. The same thing can be said of the word "blond." But we would not consider the statement, "All anxious people are blond" to be accurate. For the whole science of a profession to be accurate, both its vocabulary and the accepted body of statements that use the vocabulary must be accurate. Our efforts to make the nosology of mental illness accurate are not sufficient to make this science of treatment accurate. But they are an essential preliminary step.)

(2) "Realistic usefulness" is a complicated quality that refers here to three things: (a) the complexity of events encountered with mental disorders, (b) the complexity of the vocabulary needed to classify them, and (c) the human factors that limit our accuracy in assigning vocabulary terms to real events for the purpose of classification. There are many human factors to consider, including the number of classifications people can remember and apply consistently, the amount of dissimilarity needed between situations before people can differentiate them, and so on. Because there is no point in designing a vocabulary that is either too large or too complicated for people to use consistently, we should keep these human limitations very much in mind as we refine our nosology.

(3) Consistent application implies that every person working to control mental illness has essentially the same meaning in mind when he uses a technical or nosological term. As a result, everyone describes identical situations in essentially the same terms.

Proposals To Implement the Nosology

How might we best achieve these three goals of scientific accuracy, realistic usefulness, and consistent application in the nosology of mental illness? Several kinds of operations come to mind at once. Scientific accuracy, for example, may be introduced through (1) discovering relationship among the events associated with mental illness,

and (2) assigning terms to significant groups of associated events. Realistic usefulness must be achieved through (3) selecting and arranging terms in a way that is consistent with human limitations in classification. Consistent application requires three additional steps: (4) giving all mental health professionals some identical education and training in the use of nosological terms, (5) monitoring the uniformity with which technical terms are used, and (6) providing continuing instruction in word usage. Let us look briefly at the kinds of issues and problems involved in each of these six steps.

1. *Discover relationships among the events associated with mental illness.*—Two groups of workers, formal researchers and clinicians, are already exploring these relationships in two strikingly different ways. Researchers pursue formal programs for finding new relationships between events by consciously seeking, recording, and organizing data about them under formal rules of logic and mathematics. They usually share these findings with the rest of the professional community through published articles. Making these findings available for the task of improving our nosology requires only the comparatively simple process of finding relevant articles and arranging them in such a way that nosologists can use them easily.

In striking contrast to the systematic discoveries of researches are the unsystematic, frequently unconscious concepts that clinicians form in the course of their daily work. Although for various reasons these concepts enjoy a dubious reputation among scientists, they constitute an extremely valuable body of discoveries. The problem is how to make them available to nosologists. One promising technique for doing this is to monitor the words used in clinical records. By this technique we can get some idea from the way words are used together in clinical records how clinicians associate the events they are reporting—whether or not they are consciously aware of their associ-

ations. The words and groups of words we are particularly interested in are those that describe clinical situations and what was done about them.

But before we can hope to get enough of this information from clinical records to improve our nosology, we must accomplish three things: (1) We must improve our procedures for keeping psychiatric records so they may be reviewed more easily. For large clinics and hospitals, this almost certainly means automated record systems. These systems have recently become the focus of increasingly intense interest and strong feelings. We need say little more about them here except that they are sufficiently unfamiliar—and therefore disquieting—to most clinicians to be completely unacceptable unless they require little or no change in clinical procedures; they are easy to use; they provide useful information that cannot be produced as well in any other way; and they are nothing short of foolproof, legally as well as technologically, in protecting patients' privacy. Also, unless these systems accomplish much more than merely providing nosological research information, nobody outside a research center will want them—nor should they. Fortunately, these systems have many other advantages, and so a number of systems have already been developed.^{2 3 4} In the coming decade they will probably become the major technique for maintaining clinical files in large psychiatric facilities. (2) We must nourish the attitude among experienced mental health clinicians that improving the treatment of mental illness is a goal worth the apprehension of having one's clinical records studied by professional colleagues. Our medical and surgical colleagues have long ago come to terms with this issue, much to the benefit of their respective arts, and they are now fostering some remarkably sophisticated systems to

perform no other function than that of monitoring clinical practice.⁵ By contrast, in the mental health professions almost the only clinical records now available for study are those of students and trainees. The experience of seasoned practitioners is lost to us, except for the contributions of a handful of writers, teachers, and supervisors—who, incidentally, communicate only their conscious concepts. (3) We must make sure that technical terms used in all clinical records are translatable from one institution to another. This is necessary so that researchers (and their computers) will not overlook significant relationships between events merely because different institutions describe them with different words. There are many ways we might make records translatable. The way most likely to succeed is to create a thesaurus of terms and their synonyms used in clinical practice and to use that thesaurus to make consistent translations of all clinical records.

2. *Assign terms to significant groups of associated events.*—The important issue here is, "What is significant?" There are two ways to decide that. One way is by saying that a group of events deserves a name if it is possible to consider these events to be parts of a single disease or process. Possibility can only be determined statistically. Under a narrower interpretation, a group of events is not significant unless it is helpful to view the events as reflections of a single entity. The ultimate test of a term's helpfulness is whether or not that term improves our ability to choose the best treatments or to devise the best strategies of prevention. This important distinction between possibility and helpfulness can be illustrated by the many words some Eskimo groups have for our single word "snow." This is because it is both possible and helpful for Eskimos to distinguish between different kinds of snow. In our temperate climate it is possible but not helpful to

² The Institute of Living, 200 Retreat Ave., Hartford, Conn.

³ Reiss-Davis Clinic for Child Guidance, 9760 West Pico Blvd., Los Angeles, Calif.

⁴ Department of Psychiatry, Tulane University School of Medicine, 1430 Tulane Ave., New Orleans, La.

⁵ Commission on Professional and Hospital Activities, First National Building, Ann Arbor, Mich.

make these distinctions. We should limit our nosology and technical vocabulary to helpful terms.

Both nosological terms and other technical terms should be reviewed and updated together for several reasons. For one thing, non-nosological terms are the ones most used in teaching about mental illness. Also, technical terms are at least as important as nosological ones in communicating about patients. In fact, the diagnosis assigned to a psychiatric patient is usually much less informative than the other terms used to describe him. Finally, as I mentioned earlier, modern systems for monitoring clinical operations, including nosological studies, require some uniformity not only in nosology but also in general technical vocabulary.

In assigning names to events it is important also to record both the definition and the experimental justification for each major term. The definition should be extensive and should employ the non-nosological terms properly associated with that term. The experimental justification should cite the data supporting the use of each term. With this documentation, disagreements over word usage may be argued on the basis of facts rather than in the hazy realm of impressions and preferences.

3. *Select and arrange terms in a way that is consistent with human limitations in classification.*—This step must be guided by basic research to determine exactly what these limitations are in classifying psychiatric situations. How many classification terms can people remember and apply consistently? How does this number change when one increases the distinctiveness of the situations to be classified? How distinctive are the situations our technical vocabulary refers to? Should all varieties of mental illness be given distinctive names, or should they be named as variants of a few broad categories, such as “mild illness,” “moderate illness,” and “serious illness”? If broad categories are used, then additional descriptive terms will become more significant in the

diagnostic process. What should these descriptive terms indicate: the degree of social impairment? the quality of reality-testing? the severity of the precipitating stress?

Questions like these can only be answered after a careful, extensive program of human factors research. It should be done on an ongoing basis, both to test the usefulness of existing nosologies and to make sure that new ones are designed with their human users in mind.

4. *Give all mental health professionals some identical education and training in the use of nosological terms.*—The meaning of every term has two components for the person who uses it, a verbal one and a sensory one. The verbal component consists of the words and statements one has learned to associate with the term; the sensory component consists of the stimuli one associates with the term. Ideally there should be complete agreement among all mental health professionals on both components of meaning. The fact that we generally learn the verbal component of terms during our education and the sensory component during training suggests at least one straightforward approach to this goal: As much as possible, expose people of all mental health disciplines to the same words during their education and to the same stimuli during training.

Providing identical exposure is impossible, of course, but it is quite feasible to provide all professionals with a large body of common exposure. Common verbal exposure, for example, has almost been achieved already in that a comparatively small number of good textbooks and articles on each subject (in psychiatry, at least) tend to be used widely throughout the country.⁶ Also, there seems to be a growing internal consistency in the mental health literature. Many differences that once flared as irreconcilable issues have recently come to be

⁶ “Basic Psychiatric Literature as Determined From the Recommended Reading Lists of Residency Training Programs,” by Joan B. Woods, M.D.; Samuel J. L. Pieper, Jr., M.D.; and Shervert H. Frazier, M.D.; Baylor University College of Medicine, Houston, Tex.

viewed as tolerable differences in viewpoint that are not necessarily contradictory.

Unfortunately not nearly enough has been done yet to provide trainees with identical sensory experiences. We need an extensive library of clinical material recorded on film or television tape available for distribution—either by mail or by closed-circuit broadcasts—to all training centers throughout the country. These recordings can take many different forms, including demonstrations of particular concepts and diseases, edited vignettes from continuous treatment cases, and presentations not just of selected patients but of their families and environments as well.

5. *Monitor the uniformity with which technical terms are used.*—There are two reasons for doing this. The first is to identify terms that are often used either incorrectly or interchangeably with other terms. This information can be used in revising our nosology and technical vocabulary. The other purpose for monitoring word usage is to find institutions and individuals who use terms in a markedly idiosyncratic way. This information can be used in programs for alerting clinicians to their own idiosyncratic word usage. (See below.)

Uniformity in word usage implies that everyone who uses a term has the same meaning for it. Because there are two components of meaning, there are really two monitoring jobs to be done here, one for the sensory component of meaning and one for the verbal component.

To monitor the uniformity of sensory meaning we should see how much similarity there is in the stimuli professionals associate with each technical word. The only way to do this is to see how much similarity there is in the words professionals use to describe the same set of stimuli. Audiovisual recordings of clinical material can provide these stimuli nicely, particularly because they can be used again and again on a nationwide basis. This monitoring can be done in many settings, such as

graduate school tests and examinations for speciality board certification, where both monitoring and testing functions can be performed at the same time. The only drawback is that these settings monitor only the people who are entering the profession. Patient case registers might be more suitable to monitor experienced professionals without involving them in special testing situations. These registers record patients' involvement with all mental health practitioners and facilities in a given geographic area. Because they always note the patient's diagnosis, case registers demonstrate what nosological terms different practitioners use to describe the same patients. As a method for finding individuals who use terms idiosyncratically, this system has many limitations: One is that patients simply look different at different times and under different circumstances; the other is that some clinicians may treat many patients who appear more than once on a case register while other clinicians may treat none or very few. This approach can be useful, however, to identify terms that tend to be used interchangeably.

It is much easier to monitor the verbal meaning of terms. This can be done in two ways: By seeing what words professionals use in response to other words (for example, a diagnosis given in response to published case histories) or in association with other words (diagnoses and case histories written in the same clinical records, for example). Studying clinicians' responses to other words requires an artificial testing situation because few professionals make such responses in the normal course of their work. Studying how words are associated with each other normally is much easier to do because this is exactly what practitioners record in patient records. Probably the best way to monitor this aspect of word usage, then, is to study the records clinicians prepare in the course of their work.

6. *Provide continuing instruction in word usage.*—Programs to provide continuing in-

struction in word usage should perform at least three functions:

a. They should provide post-graduate professionals with continuing exposure to a common body of verbal and sensory information that reinforces a single vocabulary.

b. They should provide special programs of reinstruction for all professionals whenever changes are made in the nosology of mental illness.

c. They should provide remedial instruction for institutions and individuals who are using terms idiosyncratically.

Most important in planning all three functions is that they be as interesting, convenient, and painless as possible. Mental health professionals have too little free time to attend tedious programs that do nothing but polish up their vocabulary. They also tend to be uncomfortable about resuming the role of students. Consequently, all three functions should be integrated into other educational programs that mental health professionals already welcome as valuable.

The need for convenience and attractiveness is especially true in programs for alerting clinicians to their own idiosyncratic word usage. Among those who use terms idiosyncratically, it is also important to distinguish between those who are simply misinformed and those who are unusually perceptive. Misinformed clinicians who, for example, consistently assign the diagnosis of "schizophrenic reaction, chronic undifferentiated type" to patients whom their colleagues would view as having personality pattern disturbances, should be shown what their errors are and how to correct them. Unusually perceptive clinicians (and their students), who are assigning new terms to what appear to be genuine clinical entities, should be encouraged to test their findings as objectively as possible and to share them with the rest of the professional community. The advantages we gain by using words uniformly should not be bought at the price of stifled creativity.

Administrative Recommendations

All the recommendations made in the previous section can be thought of as components of one single system. In this system, information gleaned from clinical practice and formal research is collected, organized, refined, and converted by a series of operations to produce an accurate, useful nosology. The refined nosology is then communicated to the people who will apply it and, in so doing, help refine it even more. We can illustrate this system and the relationship among its components in a single diagram. (See fig. 1.) The darkened boxes indicate the information used and produced by each kind of organization playing a role in this system.

In terms of buildings, technical personnel, and equipment, surprisingly little is needed to put into action the system illustrated here. There are already many organizations that can perform its component functions. For example, a number of psychiatric facilities have begun using automated clinical record systems and many more will soon do so throughout the country. Psychiatric case registers, because of their value in carrying out demographic research and coordinating large-scale programs of treatment, have also been established in several parts of the country.^{7 8 9} and more are being planned.¹⁰ There are already two computerized clearinghouses for processing formal research information: One, the National Clearinghouse for Mental Health Information (of the National Institute of Mental Health)¹¹ is in its final

⁷ Maryland Psychiatric Case Register, Outpatient Studies Section, Biometrics Branch, National Institute of Mental Health, 5454 Wisconsin Ave., Bethesda, Md.

⁸ Monroe County Psychiatric Case Register, Department of Psychiatry, University of Rochester School of Medicine, 260 Crittenden Blvd., Rochester, N.Y.

⁹ Tri-County Psychiatric Case Register, North Carolina Department of Mental Health, 2100-C Hillsboro St., Raleigh, N.C.

¹⁰ Illinois State Department of Mental Health, 160 North La Salle St., Room 1500, Chicago, Ill.

¹¹ National Clearinghouse for Mental Health Information, National Institute of Mental Health, 5454 Wisconsin Ave., Bethesda, Md.

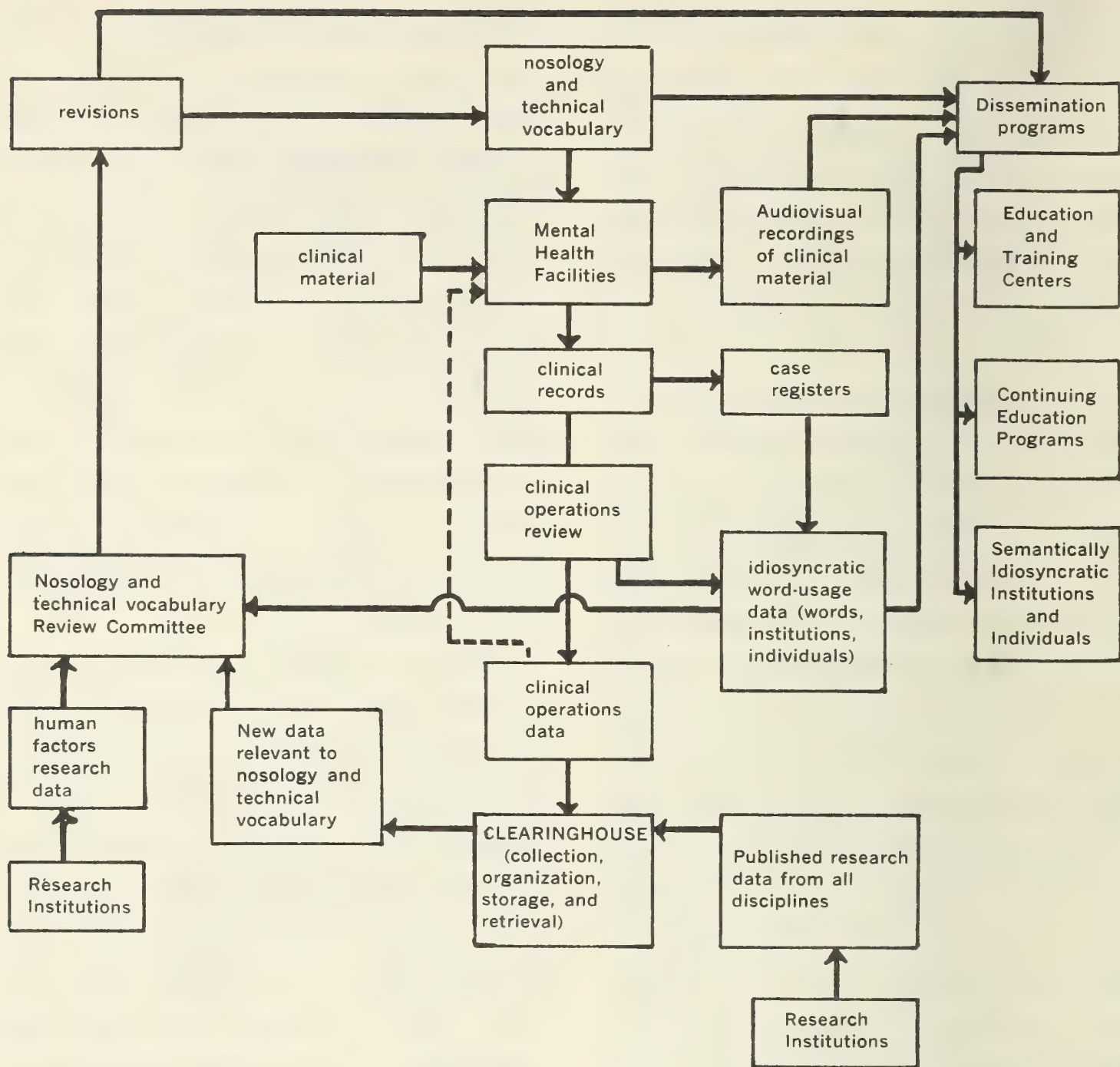


Figure 1.

development phase and may soon play a central role in refining our nosology. The other, the National Library of Medicine,¹² is already the major clearinghouse for medical information in this country and is expanding its coverage of the behavioral science literature. Within the private, academic, and governmental communities there are many organizations that can prepare excellent filmed recordings of clinical material for professional trainees and practitioners. Many professions are eager for their practitioners to become aware of new discoveries and techniques, and this has led to the development of many continuing education programs throughout the coun-

try. One could hardly wish for administrative machinery better suited to conveying the information practitioners need to use technical and nosological terms consistently. Even the Nosology and Technical Vocabulary Review Committee proposed here already has a counterpart in an active committee of the American Psychiatric Association. Coordination is the ingredient needed most if existing organizations are to generate in their present operations the information necessary to put this system into action.

To begin such a program of coordination, I would recommend the establishment of a task force having the same general composition and quality of the 1965 Nosology Conference but much smaller in size. Meeting

¹² National Library of Medicine, 8600 Rockville Pike, Bethesda, Md.

in a setting that provides leisure, informality, and freedom from distraction, the task force should begin planning with a structured agenda that includes these tasks: (1) to create a clear conceptual model of the issues involved in maintaining a nosology and technical vocabulary of mental illness, (2) to determine what data are required to administer such a model, (3) to find the organizations that can produce and process this data, and (4) to enlist the cooperation of these organizations in this nosological work.

A trial operation period should follow in which each organization performs its respective function and the task force itself does the work of the Nosology and Technical Vocabulary Review Committee. This will give each component organization a chance to work out administrative problems that could not be predicted. Having discovered from first-hand experience what the review committee members must do, the task force will then be able to enlist the right people to man the committee (preferably for a limited time in staggered rotation) and supervise their initial work before dissolving itself.

Summary and Conclusion

Improving a psychiatric nosology demands much more than just renaming psychiatric illnesses from time to time. It should be an ongoing, systematic process intimately tied to our growing understanding of mental illness. The names not only of psychiatric disorders but of all psychiatric phenomena should be reviewed periodically and revised in the light of new knowledge. This knowledge should come not only from formal research programs reported in the literature but also from careful study of the concepts formed by practicing clinicians. Each time the nosology is revised a report should be prepared explaining why each change was made and on what new knowledge that change was based.

Just as important as reviewing and modifying the nosology is the task of applying it correctly. Deliberate, vigorous steps must be taken to insure that all terms are used as uniformly as possible. One such step should involve seeing that all mental health clinicians have a common background not only of verbal instruction but of sensory experience as well. This should be done both during their initial education and training and periodically throughout their professional careers.

The administrative machinery needed to carry out this triple mission of monitoring research, updating the nosology, and encouraging its uniform application, should use existing facilities as much as possible. We can plan the exact nature and composition of this network only after studying carefully the functions it will perform. The initial Nosology Conference held in Washington in 1965 clarified many theoretical aspects of this function. Needed now is a concerted effort by a carefully chosen task force to convert these theoretical issues into a plan for action and to make this plan an administrative reality.

In many ways, establishing a program to improve a nosological system seems to be an incredible luxury. What we have outlined here will be expensive, both in money and in professional effort, even after the planning and development stages have been completed. But in considering this investment, as with any other investment, we should look also at the value to the investor. In this case, the value comes in solving many problems in research, education, training, and clinical practice that arise from having an outmoded nosology. It also lies in the role this program could play in catalyzing the synthesis of formal research and clinical experience into a powerful new resource not merely for classifying the ravages of mental illness but for controlling them as well.

The Role of Classification in Psychoanalytic Practice

*Harry I. Weinstock, M.D.*¹

The problem of the relationship of diagnosis to therapy is no new one in the history of psychoanalysis. As long ago as 1913, Freud, when chided about the lack of statistical evidence for the effectiveness of psychoanalysis, is reported to have said, "To compile statistics is at present impossible. To begin with, we work with much smaller numbers than most other doctors who devote so much less time to individuals. Then, the necessary uniformity is lacking which alone can form a basis for any statistics. Should we really count together apples, peas, nuts? What do we call a severe case? Moreover, technique changes, and what about the numerous partly analyzed cases, and those for whom treatment had to be discontinued for external reasons?"

If Freud were alive now, his reply would probably be much the same—perhaps even more pessimistic. He would find that we work with even smaller numbers of patients—an average analysis now goes on for some 4 years instead of for a few to many months. Then, he would note that at the psychoanalytic institutes there was little interest in classification, and with the widening application of psychoanalysis to all kinds of disorders well beyond the so-called transference neuroses, attention in the training institutes had turned toward the subject of "analyzability" rather than toward diagnosis. Finally, he would find that there was a great gulf between the practice of psychoanalytical institute clinics in the selection of patients and the practice of analysts in their private offices, where the vast majority of analyses are conducted. It is the policy, for example, in some institute clinics, for each applicant for psychoanalytic treat-

ment to be interviewed independently by three staff analysts, who then meet, compare notes, and finally decide whether the applicant should be accepted for analysis. Certainly, these staff members do not follow the same practice privately, nor is it likely that a private patient would find much comfort or confidence in such a procedure.

The whole question of statistics in psychoanalysis has been studied by a statistician, J. B. Chassan, and reported in two articles, one in 1953 (1) on the "Role of Statistics and Psychoanalysis" and the other in 1956 (2) on "Probability Theory and Psychoanalytic Research." He concludes his first article as follows: "Finally, that it is impossible to be absolutely certain that perfect comparability can ever be achieved with data as complex as that relating to psychoanalysis, does not imply that a striving toward such an objective will not exert ultimately a positive influence on psychoanalytic practice and theory. While it is quite likely that statistics can help provide answers to some psychoanalytic questions in a comparatively short time, a realistic view requires that optimum benefit can be achieved only after many years of careful collection and study of psychoanalytic data."

In Knight's 1941 (3) review of previous psychoanalytic statistics, he gathered nearly 1,000 cases from the reports of the Berlin, London, and Chicago Psychoanalytic Institutes, and included those from the Topeka Institute. These 1,000 cases dated back to 1920 and extended up to 1941. In most instances the studies were retrospective, i.e., studies of past records, not originally planned for research. Various classifications were used, and in that period many developments in theory, technique, and length of treatment took place. His compilation grossly suggested that 63 percent were "apparently cured" or "much improved" in the psychoneuroses, 50 percent similarly in the sexual disorders, 55 percent in the character disorders, 78 percent in the organ

¹ Chief, Department of Psychiatry, Mount Sinai Hospital Services, City Hospital Center at Elmhurst, Elmhurst, N.Y.

neuroses and psychosomatic conditions, and 25 percent in the psychoses.

Since these composite figures lack any uniformity of classification or mode of evaluation, they are of little, if any, value—except to indicate that there has long been some interest in statistics among psychoanalysts.

The American Psychoanalytic Association has officially been supporting a program of factfinding almost continually since 1947. At that time, a committee on “The Evaluation of Psychoanalytic Therapy” was set up under the chairmanship first of Dr. Jules Eisenbud, and then under Dr. Jean Cushing. The minutes, correspondence, and preliminary studies of these committees, filled about 100 typed pages, but included little participation by the members of the association. It concluded its final report in 1952 with the following statement: “If a matter of many years, perhaps decades, is of no consequence, then the slow compilation of statistics and gradual erosion of resistances on the part of the membership will be the method of choice. In any event, it is the conviction of a number of members of this committee that the evaluation of psychoanalytic therapy is an extremely valuable and necessary project and must be continued.”

After an inactive interval of about a year, and at the instigation of Dr. Knight, then the president of the association (1952–54), the membership voted to authorize another committee “with more modest initial aims * * * to establish a method and tradition for pooling certain data of psychoanalytic practice.” If this were successful, a more comprehensive program of statistical research might then be developed. It can be said that to some degree a method was established, and also a tradition, but the diagnostic problem, recognized early in the study, led to its abandonment after about 4 years of operation.

Asked to chair this committee, called the central fact-gathering committee, I reserved decision until I hit upon a method of collecting data that would overcome some

of the reality-based resistance, i.e. the fear of exposure of patients or analysts themselves. Consulting with the IBM Corp., it was learned that their service bureau performed various highly confidential activities for industry that could easily be applied to this study. The method was simply to use the machinery of their service bureau as an intermediary that would give each analyst-participant a code number and, via the machine, make it possible to communicate with the corresponding participant whenever necessary. The analyst, in turn, would give each of his patients then in treatment a number beginning with 1 and going up to a total of 25 for the long-term plan of the study. Thus, only the analyst himself knew the identity corresponding to the code number (1–25) of his patient, only the confidential service bureau could match the code number of the analyst to his identity card, and yet the committee could communicate through the intermediary of the service bureau to the analyst for corrections or additions to any of his questionnaires, a carbon copy of which remained with him.

A description of this method was then reported to a meeting of the association at which occasion about three-quarters of the members present agreed by a show of hands to participate in the study. Having thus overcome the fear of exposure, the committee then set out to formulate the questionnaire. The help of the membership was sought in every way in developing this, but little aid came forth. Perhaps the committee should have taken this as an indication of the ultimate failure, but it was hopeful and went on as best it could to develop its program.

After completing the relatively factual items of the questionnaire, the problem of diagnosis was considered. The APA Manual having recently been issued, it was decided to use this, but in a more complete manner than it itself recommended. Participants were instructed to familiarize themselves with the manual, the relevant portion of which was sent gratis to each. They were

told to write in the APA number of the primary or general psychiatric diagnosis first, i.e. the basic, all-inclusive personality type or underlying psychiatric disorder and then the subtype of this beneath it, using a descriptive word or two as required. The secondary diagnosis would then consist of the symptomatic diagnosis and, likewise, the subtype, by another descriptive word or two, as the following examples show:

A case of fetishism, shoe-fetishism, for instance, in a passive-aggressive personality, passive-dependent type would be listed as:

Primary diagnosis:

000-x52 (i.e. passive-aggressive personality)

Subtype passive-dependent (write in)

Secondary diagnosis:

000-x63 (i.e. sexual deviation)

Subtype fetishism-shoe (write in)

A case of phobia of dogs in a schizoid personality would be listed as:

Primary diagnosis:

000-x42 (i.e. schizoid personality)

Subtype

Secondary diagnosis:

000-x04 (i.e. phobic reaction)

Subtype dogs (write in)

A case of asthma in a compulsive personality would be listed as:

Primary diagnosis:

000-x53 (i.e. compulsive personality)

Subtype

Secondary diagnosis:

003-580 (i.e. psychophysiological respiratory system)

Subtype asthma (write in)

A case of homosexuality with pedophilia in a paranoid personality would be listed as:

Primary diagnosis:

000-x44 (i.e. paranoid personality)

Subtype—

Secondary diagnosis:

000-x63 (i.e. sexual deviation)

Subtype pedophilia (write in)

In all instances, the subtypes were then given code numbers, the total of these com-

ing to about 300. As a consequence, we have a large number of subtypes of disorders of the numbered diagnoses. For example, there are listed under dissociative reaction 11 different disorders, e.g. fugues, depersonalization, dream states, epileptoid reaction, rage reaction, stupor and hallucinations, emotional inhibition, and freezing. Under conversion reaction, there are listed some 30 different types—e.g. convulsions, vaginal pruritus, impotence, torticollis, sexual anesthesia, tremor, tic, frigidity, amnesia, tetany, deafness, stuttering, cough, blindness, ejaculation retardata, dysphagia, ataxia, etc. Under phobic reaction there are some single and multiple phobias including fear of buttons, of work, of marriage, of being alone, of being with people, of disease, of traveling, of choking, of dying, of insanity, of eating, of men, of women, of germs, etc. Under obsessive-compulsive reaction, there were at least 24 types, e.g. compulsive doubting, ceremonials, hand-washing, aggressive impulses, dirt, guilt, masturbation, stealing, crying, blinking, etc. Under personality trait disturbance, other—i.e. a group that could not apparently be included under the passive-aggressive, compulsive, or emotionally unstable personality, there were some 30 types of disorders that would in many instances probably be described by the analyst, as character neuroses, e.g. pseudo-passivity, masochistic personality, moral masochism; inhibited personality, work inhibition, fate neurosis, frigidity, impotence, promiscuity, prostitution, Don Juanism, phallic character, masculinity complex, latent homosexuality, potential homosexuality, homosexual phantasies, recurrent bad choice of heterosexual partner, adolescent exaggeration of manliness; temper outburst, explosive outburst, acting out; gambling, kleptomaniac; infantile, oral dependent, narcissistic, schizoid, hysterical, epileptic-aggressive, oral aggressive, etc.

One might note that all kinds of diseases were listed as “psychophysiological re-

actions," even, for example, psoriasis, alopecia areata, acne, rheumatoid arthritis, recurrent respiratory infections, Raynaud's syndrome, retinal vascular spasms, anemia, ileitis, mucous colitis, benign paroxysmal peritonitis, renal colic, diabetes, sterility, Parkinson's disease, epilepsy, etc.

It can be readily seen that in spite of this attempt at a fairly comprehensive diagnostic system that might approach a description of the total clinical picture, there is a certain amount of overlapping, a great deal of ambiguity, and there is no certain way of ascertaining which cases are similar, much less identical. The same symptom is listed under a number of disorders. Some would question whether in many cases the particular subtype should be listed under the diagnosis given. Further confusion is created by the increasingly common use of the term "psychosomatic" or "psychophysiologic."

The questionnaire was divided into two sections—an initial report and a final report (q.v.). When these came to the service bureau, they were then forwarded to the chairman of the committee who checked each of them. Wherever clarification was needed, a prepared form was checked off and returned with the questionnaire to the service bureau, and the machine, matching the code number of the participant to his identity card, returned it by mail to him. In some instances, the same questionnaire went back and forth from three to seven times in efforts to clarify various aspects, usually of diagnosis, or to correct obvious inconsistencies.

At the time the study actually began (1953), there were about 500 members of the association of whom 348 agreed to cooperate, and about 500 senior candidates and graduates of institutes (i.e. those engaged in analytic practice, but not yet members of the association), of whom 412 agreed to cooperate, i.e. a total of 760 out of a possible 1,000. Within a year, we had initial data on 4,400 cases in treatment from some 500 participants. A study by

the service bureau showed that of then 13 institutes, the participation rate ranged from 60 to 90 percent, averaging about 65 percent. The same 65 percent participation was noted among training analysts, ranging from 50 to 100 percent.

When the figures were analyzed according to length of membership in the association, the participation ranged from 44 percent of those who had been members for over 20 years, to 70 percent of those who were members for 10 years. Thus, there was a fairly wide representation of the membership who agreed to share their data in the study.

On an early sample of over 3,000 cases, the following data may be of interest:

1. Sex:
 - 48 percent were males.
 - 52 percent were females.
2. Race: 99 percent were white.
3. Religious background:
 - 43 percent were Protestants.
 - 45 percent were Jewish.
 - 10 percent were Catholic.
 - 1 percent were uncategorized.
4. Age:
 - 26–35 years, 48 percent.
 - 36–45 years, 27 percent.
 - 46+ years, 8 percent.
 - 13–25 years, 4 percent.
5. Educational level: 60 percent were college graduates compared to the 6 percent average in the United States at that time.
6. Economic data: The patient or his family paid for treatment in 93 percent of the cases; an outside source in 4 percent, and no answer given in 2 percent.
7. Income:
 - 40 percent had an income over \$10,000.
 - 57 percent had an income under \$10,000.
8. Previous psychoanalysis: 15 percent had had psychoanalytic treatment.
9. Previous psychotherapy: 35 percent had had previous psychotherapy.

10. Present treatment:
54 percent were checked as being in analysis and 46 percent in psychotherapy.
From 60 to 65 percent of cases in the neurotic group were in analysis, about 40 percent in the borderline group, and 20 percent in the psychotic group.
11. Use of psychological projective tests:
These were done in about 25 percent of cases and were checked as agreeing with the clinical diagnosis in 75 percent of these.
12. The frequency of the diagnostic groups:

	<i>Percent of total</i>
Neurotic reactions	39
Personality trait disturbances (character neuroses)	33
Personality pattern disturbances (borderline cases)	11
Psychotic reactions	9
Perversions	5
Addictions	2

13. The frequency of the specific diagnoses:

	<i>Percent of total</i>
Passive-aggressive personality..	16
Compulsive personality	14
Anxiety reaction	14
Depressive reaction	9
Schizophrenia	7
Phobic reaction	6
Obsessive-compulsive reaction	6
Schizoid personality	6
Perversions	5
Conversion reaction	3
Emotionally unstable personality	3
Dissociative reaction	2
Inadequate personality	2
Cyclothymic personality	2
Paranoid personality	2
Addictions	2
Paranoia and manic depressive reactions (each)	1

In accord with the general impression, dissociative, conversion, depressive, and phobic reactions were much more frequent in females than in males, while obsessive-compulsive reactions were diagnosed somewhat more frequently in males than in females. On the other hand, anxiety reaction was diagnosed with equal frequency in both sexes, as was the schizoid and paranoid personality and schizophrenia. Sexual deviations were reported two to three times more frequently in males than in females.

In their order of frequency, the chief presenting symptoms checked were interpersonal difficulties, anxiety states, non-psychotic depressions, work inhibition, neurotic sexual disorders, compulsions and obsessions, phobias, psychosomatic disorders, perversions, and psychoses.

In the last few years, a new committee has been reviewing some of the data with the aid of the computer and of statisticians. This committee abandoned the diagnostic data in the questionnaires altogether and is making a study of presenting symptom syndromes. By clustering these presenting syndromes, it has established nine categories of cases: the schizophrenic, psychosomatic, paranoiac, sexual deviate, hysteric, compulsive, anxiety, depression, and anxiety and depression, leaving one group of cases uncategorized. An analysis of analytic experience with a large group of these categories will soon be published.

In the second year of the study, we issued a preliminary report on selected data from 6,500 initial questionnaires and 1,000 final questionnaires on these cases, many of which represent cases, it must be recalled, that had already been in treatment for varying lengths of time.

We found that in the first 100 cases each, of the cases diagnosed anxiety reactions, passive-aggressive personality, and compulsive personality, and an unrelated group of 838 cases of all kinds of disorders, just about one-third of each had been checked

as having completed their analysis or psychotherapy. The average duration of analytic treatment in completed cases was 3 to 4 years, at 4 to 3 hours' weekly frequency, and for psychotherapy 1½ years at 1 to 2 hours' weekly frequency.

Very soon after the study began, we were faced with the problem of further clarification of diagnosis. This perhaps was responsible for the failure of 20 percent of those who had initially agreed to participate to send in any questionnaires at all. Through the intermediary of the IBM, about one-fourth of all the questionnaires that did come in were returned for completion, correction, or amplification, and we received innumerable complaints about the difficulties of compressing the diagnosis into our expanded APA diagnostic system. At the meetings of the association, we heard again and again from participants who said that they had discontinued or would discontinue their cooperation because of this. Others stated that they were using just one diagnosis, most often passive-aggressive personality, or compulsive personality, for most of their cases, while some reported that they called almost all of their cases pseudoneurotic schizophrenia.

As another example of the diagnostic problem, some material prepared for a panel on "Phobias and Their Vicissitudes" is of interest. In a random sample of 230 cases in which phobias were reported in the diagnoses, one-half of these listed phobia as a primary diagnosis, with the additional mention of the type of phobia. In the other half, that is, in about 120 cases, in addition to the diagnosis of phobia, the following were listed either as primary or secondary diagnoses:

	Cases
Anxiety reaction	26
Dissociative reaction	2
Conversion reaction	2
Obsessive-compulsive reaction....	11
Depressive reaction	10

	Cases
Paranoid, schizoid, or cyclothymic personality (borderline disorders)	12
Personality trait disturbance (character neuroses)	32
Sexual deviation	5
Drug addiction	5
Psychophysiologic disorders.....	16

There is no way of knowing whether such additional diagnoses should also have been included in the group listed only as phobia. Whether it was the prominence of the symptom, the viewpoint of the analyst regarding diagnosis altogether, his clinical evaluation of the total personality structure or disorder, his concept of schizophrenia, any or all of these probably play a role from case to case, in listing more or less complete or incomplete diagnosis.

Further evidence of this difficulty may be seen in a later questionnaire study of 61 cases of phobias that had been reported as completely analyzed. To the question "Is there any doubt in your mind regarding the diagnosis listed?" 14 of the 46 who replied to the questionnaire replied in the affirmative and submitted alternative diagnoses such as anxiety reaction, mixed neuroses, latent psychoses, schizophrenia, dissociative reaction, and depressive, compulsive, and passive-dependent character. To the question, "Did you think at any time that an underlying psychosis was present?" 8 of the 46 replied in the affirmative.

In another random sample of 425 cases that included phobia in the comprehensive diagnosis, 60 percent were in analysis, and 40 percent in psychotherapy. Of this sample, 7 percent had been in psychiatric hospitals, 47 percent had had previous psychotherapy, and 12 percent had had previous analysis. In a very high percentage of cases, the symptoms had been present for more than 2 years before treatment was started. In only three instances had the symptoms been present as short a time as

3 to 6 months before therapy. How difficult it is for a single analyst to gather data about one disorder is evidenced by the report of one of the analysts at the panel, that in 20 years of analytic practice, he had treated but four cases of phobias. The two other panelists, with about as many years of experience, had treated no more between them.

It should be mentioned that there is an increasing number of analysts who are beginning to question the diagnosis of neurosis altogether. As the psychologic concept of schizophrenia is being widened, more and more analysts are privately beginning to state their belief that all the symptomatic disorders are cases of schizophrenia; that it is not a question any more of whether it is schizophrenia, but what kind of schizophrenia and how severe. In support of this to some degree is the follow-up of some of the so-called classical neurotic cases analyzed by Freud, a number of which undoubtedly would now be classified as schizophrenia.

As the members of the committee were finding the same difficulties in listing their own cases, it was decided to invite the membership to form groups at the various institutes to study the problem of diagnosis, but as far as could be determined, only one group met and that for only one discussion. Following this, it was decided by the committee to seek foundation support for two purposes—the study of diagnosis, and the study of methods for the evaluation of results. A plan was prepared for quarterly meetings of committees of interested members to work on these two separate problems, with the assistance of members of the disciplines such as are assembled in this group, but the grants not being given, hope for the continuation or the success of this venture was gradually abandoned. Fewer and fewer reports came in, and the study was discontinued in 1958. A final summary of the data was issued to the membership and the limitations of its

validity stressed, especially in the diagnostic and evaluative sections. Because of these limitations, it was decided not to publish the data, as this would only create meaningless controversy.

During the latter part of this study, some related questionnaires were submitted to the membership of the association and are herewith reported as suggestive of the varying attitudes among analysts on the question of diagnosis and therapy.

In 306 cases of neuroses, that were reported as completely analyzed, a supplementary questionnaire for additional information brought forth the following from 210 replies:

- (1) In 28 percent of these there was doubt in the analyst's mind regarding the diagnosis.
- (2) In 25 percent the analyst stated that at some time during the treatment he thought there was an underlying psychosis present.

In order to gather the opinions of members of the association regarding the expectation of results from analytic treatment of various disorders, the following questionnaire was circulated. "Given a young person whom you could analyze for 4 years or more, with all conditions favorable, what would your expectancy of result be, in percentage, of cure, improvement, and failure, for a list of neuroses, personality trait disturbances, schizophrenia, schizoid personality, and homosexuality?"

From the then about 700 members of the association, we received only 120 replies with great degrees of variance, e.g.

- (1) 45 percent expected no cure in any of the conditions.
- (2) 35 percent expected from 50 to 100 percent cure in the neuroses less progressively in the personality disturbances, psychoses, and perversions.
- (3) 20 percent expected 5 to 40 percent cure in the neuroses, less in the other conditions.

Of those 45 percent who expected no cure or a low percentage of cure, 50 percent expected moderate improvement, 45 percent great improvement, and 5 percent slight improvement.

Of those 35 percent that expected some percentage of cures, the average expectancy of cure was:

- (1) 50 percent in anxiety, conversion, and phobic reactions.
- (2) 33 percent in dissociative, obsessional, and depressive reactions.
- (3) 10 percent in schizophrenia.
- (4) 20 percent in homosexuality.

If these figures are combined with the reports of those expecting no cures at all, the overall expectancy of cure, from this limited sample of the membership, would be approximately half of the percentage stated.

Limited and gross as this questionnaire was, and even though representing only about 17 percent of the members, it shows clearly that there is a wide divergence of opinion regarding the efficacy of psychoanalytic treatment in various conditions, that though analytic treatment is employed in a broad range of disorders, the expectancy by analysts of good results is higher in the transference neurotic disorders than in the psychoses and perversions.

The following remarks were pencilled in on three of those questionnaires:

- (1) There is a meaningless distinction between the neurotic reactions and the personality trait disturbance.
- (2) The motivation of the patient is more important than the diagnosis. Therefore, there is no point in filling out the questionnaire.
- (3) I would not judge prognosis in terms of diagnosis.

Seeking help and suggestions from the membership, we issued another questionnaire, again receiving replies from about 15 percent of the membership. Many encouraged the continuation of the study, but

many reported their difficulties with diagnosis. The following are some of the complaints received

- (1) When carrying a case to completion, we meet a whole gamut of psychopathological potentialities, with differences of relative weight of psychopathological factors making it all difficult to put into classification.
- (2) Change the diagnostic procedures radically—from a number of responders.
- (3) The study makes perhaps the unexpected contribution of pointing up the fact that the APA nomenclature is not used by most analysts in their own thinking, even when thinking clearly. There is something unreal about going through the motions of using it here.
- (4) This program is seriously handicapped by the assumption that psychiatric diagnoses form a workable system of classification and that these diagnoses apply to the psychoanalytic situation and psychoanalytic treatment. This implies a high degree of similarity between traditional medical models of disease and the sort of problem dealt with by psychoanalysts, an altogether untenable position. Cases cannot be fitted into the diagnostic categories.
- (5) What are the indications of a completely analyzed person? Intensive studies should be made of cases with good therapeutic results. One might find psychotherapy other than analysis was producing better results.

From time to time, members raised questions regarding the reason for analysis rather than psychotherapy in cases with the same diagnosis. This led to another brief questionnaire answered again by only 15 percent of the membership. Members were asked to list the number of patients they

had in psychotherapy, checking off the following reasons for choosing psychotherapy rather than analysis:

- (1) Considered too mild for analysis.
- (2) Considered too severe, not likely to benefit from analysis.
- (3) Economic, unable to afford analysis, though analysis was method of choice.
- (4) Patient has been analyzed and returns for occasional psychotherapy.
- (5) Patient was being prepared for analysis.
- (6) Other reasons.

The summary of this study indicated that in the practice of the 100 analysts who replied, there was an average of one-half a case in therapy because it was too mild for analysis, two cases because they were too severe for analysis, 1½ cases that were in therapy rather than analysis for economic reasons, one case for therapy after analysis, and one case in preparation for analysis. There was an average of one-half to one case per analyst in therapy rather than analysis for a variety of other reasons—e.g. too young, too old, too rigid, distance, etc.

Again, even though the sample is small—about 15 percent of the members—it is evident that other factors than diagnosis play roles in the choice of treatment. Parenthetically, this is sometimes the case in physical disorders and their treatment. Sometimes the specific therapy for a condition may be fatal for a specific individual on account of his special sensitivity, etc. Age, general condition, and factors in addition to, and other than, the specific pathologic disorder often determine both the selection of medical or surgical procedures and the time for instituting them. However, it cannot be denied that there is a far higher degree of specificity of diagnosis and therapeutic procedure in physical disorders, particularly those that are acute. Whether this is equally true in chronic physical disorders is a moot question. Need-

less to say, most psychiatric disorders are chronic. They have had early and long roots and often multiple etiologies and have been influenced by many long-lasting factors. Perhaps in all chronic disorders, whether physical or psychological, the entire concept of treatment is realistically considered from the viewpoint of the availability of therapies and the justifiable limitations in goal expectations, rather than from the diagnoses alone. Further, it will be interesting to hear from the statisticians whether it is at all possible to establish statistical norms for the chronic lifetime disorders that we deal with. Attempts to establish such patterns for heart disease have been met by diagnostic difficulties even in this area that is more readily subject to measurement and classification. It was learned from the American Heart Association that in spite of setting up clear descriptions and definition for various disorders of the heart for statistical purposes, the expert in heart disease often found them inadequate, while the inexperienced found these definitions too demanding. One may also question whether it is at all possible for a large group to cooperate in any study of diagnosis and therapy because of the great variance in skills, experience, and viewpoints that tend to be operative whenever more than a few engage in any study. Whether foolproof questionnaires can be devised in the psychiatric field remains to be seen.

Some years ago, at one of the psychoanalytic institutes, a series of seminars was held over many months on the subject of the "Indications for Psychoanalysis." Its purpose was to spell out the content of those deliberations about analyzing which go into the analyst's initial assessment of each patient. Reviewing the bulk of clinical psychoanalytical practice, i.e. neuroses, character disorders, mild perversions, mild impulse disorders, mild borderline states, etc., it was forced to the position that the lack of nosological clarity, as well as the

impossibility of making any meaningful quantitative gradations within these diagnostic concepts, compelled abandonment of the diagnostic approach in favor of a more dynamic and less descriptive orientation.

Thus, evaluation of the functioning of the ego, in terms of the analyst's conception of ego function, was considered as primary in the question of analyzability.

Some of the ego functions of consequence in this regard were considered to be tolerance of frustrations, tolerance of the passive role in treatment, and tolerance of anxiety. Next in importance for analyzability was considered the capacity for introspection, which included the ability to make self-observations, self-judgments, and insight. Adequate reality-testing ability was obviously a requirement as was the ability to form transference relationships and object relationships. The analyst's opinion of the patient's capacity for free association and his judgment of the patient's motivation and of various resistances, entered into the question of analyzability.

All of those elements involve judgments, interpretations, intuitions, evaluations, and perhaps many other less-recognizable activities on the part of the analyst, and in all likelihood defy measurable description and statistical validation at this time.

Summary

Psychoanalytic therapy began before the turn of the century when modern psychiatric classification was in its infancy. Quoting from Freud in 1940 (4) "Psychoanalytic therapy was created through and for the treatment of patients permanently unfitted for life. At first my material consisted entirely of patients who had tried everything else without success and had spent long years in sanatoria. The conditions under which this method is indicated, or contraindicated, can scarcely be definitely laid down as yet."

As psychiatric classification developed, there were attempts to interpret the then-described psychiatric disorders in terms of psychoanalytic theory of personality of that time. The multiplicity of factors involved in the considerations of the applicability of treatment and the ambiguity of diagnosis still continuing, the relation of diagnosis to therapy remains a problem.

Nevertheless, attempts have been, and are continuing to be, made to clarify this problem. In this country, the American Psychoanalytic Association has officially been promoting studies of this kind.

The findings of the central fact-gathering committee, which functioned between 1954 and 1958, that are relevant to the question of the role of classification in psychoanalytic practice indicate the following:

- (1) Almost 90 percent of the patients in treatment by psychoanalysts are in the neurotic rather than psychotic diagnostic groups.
- (2) Psychoanalysis is the treatment carried out in
 - 60 percent of neurotic reactions and character neuroses (personality trait disturbances)
 - 40 percent of borderline cases (personality pattern disturbances)
 - 20 percent of schizophrenic disorders

The remainder of the cases of the same groups of disorders are in psychotherapy.

- (3) Psychoanalysts found such great difficulty in fitting their cases into the American Psychiatric Association standard classification that their cooperation with the study progressively diminished to a complete stop.
- (4) Interest in the question of analyzability has increased among analysts as both the interest in descriptive diagnosis has diminished and the scope of analysis has widened.

- (5) The views of analysts differ greatly on therapeutic effectiveness in various disorders. Judgments are not based on clinical diagnosis alone.
- (6) Factors other than clinical diagnosis play a role in determining whether a case is in analysis or in psychotherapy.
- (7) Even after a complete analysis of a group of 46 cases of phobia, there was diagnostic doubt in the mind of the analyst in nearly a third of the cases, and the possibility of psychosis considered in one-sixth of them.

Whether the same questions raised by Freud regarding diagnosis, severity, external factors, changes in technique, and many others inherent in the fact of chronicity and uniqueness of psychiatric disturbance, can be answered to an important degree by the newer knowledge of statistics and the new computer science, remains to be seen. Surely, psychoanalysts will welcome efforts to correlate theory and practice for the benefit of both.

BIBLIOGRAPHY

- (1) CHASSAN, J. B., "The Role of Statistics in Psychoanalysis," *Psychiatry* (1953) 2: 153-165.
- (2) Ibid., "On Probability Theory and Psychoanalytic Research," *Psychiatry* (1956) 1: 55-61.
- (3) KNIGHT, R. F., "Evaluation of the Results of Psychoanalytic Therapy," *American Journal of Psychiatry* (1944) 98: 434-446.
- (4) FREUD, S., *Collected Papers*, vol. I (1904), "On Psychotherapy."

The following, in addition to the author, were members of the central fact-gathering committee during various periods of its functioning: Drs. Leo H. Bartemeier, Roy R. Grinker, David Kairys, Lawrence C. Kolb, Lawrence S. Kubie, Alfred O. Ludwig, Milton L. Miller, Milton Rosenbaum, and George W. Wilson, and consulting statistician, Jack B. Chassan, Ph. D.

The full report of the work of the original fact-gathering committee and of the later committee, under the chairmanship of Dr. David Hamburg, will be published in the coming year.

Discussion

LEON EPSTEIN, M.D., *Associate Medical Director, Langley Porter Neuropsychiatric Institute, University of California Medical Center, San Francisco, Calif., Chairman*

(Papers: "The Role of Classification in Hospital Psychiatry," Henry Brill, M.D., Director, Pilgrim State Hospital, West Brentwood, N.Y. "The Role of the Classification System in Outpatient Psychiatry," Elmer A. Gardner, M.D., Director, Division of Preventive Psychiatry, Department of Psychiatry, University of Rochester. "The Role of Classification in Psychoanalytic Practice," Harry I. Weinstock, M.D., Chief, Department of Psychiatry, Mount Sinai Hospital Services, City Hospital Center at Elmhurst, Elmhurst, N.Y.)

OPEN DISCUSSION

Dr. EPSTEIN. The papers presented have been both informative and provocative. The meeting is now open for discussion.

Dr. GREENHOUSE. Dr. Weinstock, you indicated a new study committee in psychoanalysis abandoned the diagnostic data from the questionnaires and is trying to substitute something like a symptom syndrome. I wonder if you would indicate why they abandoned the diagnostic data?

Dr. WEINSTOCK. I think it is obvious why the diagnostic listings were pretty much abandoned: They were of little value because they began to be used as scrapbaskets. In the questionnaire there was a listing of symptoms, and this committee felt this might be a better index of types of disorders. (Some were left out because they comprised a rather indiscriminate group.) They have clustered these symptoms. The report will be based upon this clustering and will be published relatively soon.

Dr. LEHMANN. Were symptoms chosen rather than diagnoses because they seem to be better indicators of the underlying pathology, or better indicators as far as reliability of diagnosis is concerned, or

better indicators as far as correlation with treatment response is concerned?

Dr. WEINSTOCK. My impression is they were considered better indicators of diagnosis.

Dr. BRILL. I would like to ask Dr. Weinstock whether, in truth, a symptom as he uses it may be only an abbreviated diagnosis, leading to circular reasoning. For example, let us consider the terms "emotional inadequacy" and "*la belle indifférence*." If we think the patient is a schizophrenic we call it "emotional inadequacy," if we think she is hysteric we call it "*belle indifférence*." So when you name the symptom you already imply something about your orientation as to the syndrome that lies behind it.

Dr. WEINSTOCK. I think that is true, to a degree. But I think you have to see this in its setting. There was a long list of symptoms, I think maybe 17 symptoms, and one could check off a number of symptoms. They will try to correlate the symptom clustering with the diagnoses and see whether something is to be learned from it. I don't think this thing is settled.

Dr. GREENHOUSE. I would like to ask Dr. Gardner whether he wasn't implying the need for a multidimensional designation of a name for a patient, such as an ordered vector, where the first component would be the conventional diagnosis and the second the severity, the third, disability, and so on?

Dr. GARDNER. If I understand your question correctly, that is essentially what I implied. I speak of multidimensional systems, but perhaps this is incorrect. Although some of the parameters are dimensional, others are not.

Dr. KLERMAN. The American Heart Association has done impressive work in my opinion in developing for heart diseases exactly the kind of schema which Dr. Gardner proposes for our field. I think a similar system—would add clarity to our

communication problems. The cardiac system outlines criteria for classifying the patient along a number of different dimensions. It includes an etiologic dimension, including a category of "unknown etiology." It also includes a system for grading disability which allows for changes as the patient changes during his illness. It also allows for anatomical, radiological, and EKG diagnoses, which are well developed in the heart field. There may not be any corresponding techniques in our field, except perhaps certain diagnostic psychological tests. I believe this would be a useful model for the American Psychiatric Association classification to follow. Sometimes we talk etiologically; other times we talk in terms of clinical phenomenology; other times we talk in terms of dispositional justification. No one category can cover all the uses to which diagnoses are applied.

Dr. BAHN. A committee of the American Orthopsychiatric Association, chaired by Dr. Visotsky, has spent the last year developing a proposal for a multidimensional classification scheme that will, we hope, supplement the classification of the American Psychiatric Association. It will cover many of the points mentioned by Dr. Gardner and Dr. Klerman. I am certain that Dr. Visotsky would be glad of any interest you might have in working with the American Orthopsychiatric Association's committee. It should be emphasized that the committee is interested in psychosocial disorders of all age groups. There is need, for example, to describe the level of functioning of the aged person for purposes of Medicare.

Dr. FINK. May I ask if the APA committee considered a classification following a multidimensional model similar to that advocated by the American Heart Association?

Dr. BRILL. There is a provision which was never implemented in the APA for just this sort of thing. And hopefully this could be implemented if somebody wanted to pur-

sue it—the intensity, or the amount of disability, can be specified.

Dr. FINK. As I understand the Heart Association classification, intensity is not the only variable. Disability is only one of four elements. The additional aspects include the physiological and the etiologic. Did the APA consider the multidimensional, or multicomponent classification? And, since they apparently have not recommended it, if they considered it, why was it discarded?

Dr. BRILL. Because the statisticians turned pale. And some of them fainted at the idea of trying to handle this kind of material on a multidimensional basis. They objected bitterly that it would reduce classification to innumerable small cells if we were to introduce multiple independent axes of classification. There is a statistician who wants to respond right now.

Dr. KRAMER. It wasn't the statisticians who turned pale or fainted at the suggestion; it was concern on the part of the clinicians who have to complete a more extensive form in a more systematic way. The statistical handling of N-variables doesn't bother statisticians. They know how to handle N-variables. It's the matter of getting the clinicians who utilize various kinds of classificatory systems to utilize them systematically from time to time and place to place. This is a major problem.

Dr. WEINSTOCK. I called the Heart Association and discussed their system, because I was interested in seeing how we could set up a system that would be more valid. And do you know what they told me? "This system is all very nice. But it is not good enough for the experts, and it is too good for the inexpert." So we, too, had a lot of difficulty with our classification.

Dr. PASAMANICK. I think we ought not to confuse classification with recording. If you recall, classification is an ordering on a theoretical basis, with some abstract concept in mind; whereas, recording is merely

checklisting some items. One of the problems involved in the psychiatric use of disability, for instance, is that we already contain within the phenomenologic aspect of the classification etiologic aspects. We also have phenomenological categories that already contain disability estimates. So if we used an additional axis of classification, including disability or severity, we would be contradicting ourselves, or in danger of further confusing the issues. I think the whole problem revolves around the tremendous unreliability and lack of validity in our present classificatory system. To introduce additional axes of classification would only serve to further confuse the issue.

Dr. FLEISS. Perhaps it is because the current classification system attempts to combine all of these different axes into one dimension that unreliability and lack of validity have resulted. The alternative that has been suggested, a multidimensional, or a multicompartmentalized classification system, will consist of classifying a patient along each one of these dimensions. An example would be the classification of accidents. Hospital administrators want to know one aspect of the accident; the surgeon wants to know what bone was broken; the person concerned with prevention wants to know whether the patient was in a car accident, in a plane accident, in a bus accident, or merely a pedestrian. A number of labels, rather than a single one, are therefore to be attached to each patient.

Dr. BRILL. In talking about multiple diagnoses, is this envisioned as multiple independent diagnoses, or is it planned to work on the permutations of these diagnoses? Because the statistical problem, as I understand it, arises when one attempts to intercorrelate these things and reach some sort of a master diagnosis.

Dr. GARDNER. I would like to elaborate on what Dr. Pasamanick said. I believe there

is confusion about recording and classification. In the practice of psychiatry various combinations of our classifications are actually used. These combinations are employed to present a clinical picture and are often an abuse of the classification system. This represents a problem of recording and must be differentiated from the need for a new system. However, this does not negate the need for a new classification system. I would like to elaborate on Dr. Zubin's concept of disease.¹ I think that we do have the problem of what we mean by disease. Perhaps part of the problem in classification stems from the broadening of the concept, disease, particularly in the last 40 or 50 years, and particularly in outpatient practice. For example, in a court clinic where we may interview many delinquents and we are faced with the question: "What are you going to call this"? A disease? A defect? A disorder? Can you include it in our classification system or not? Perhaps this illustrates the need for a new classification scheme, or some new categories.

Dr. GARFIELD. There seems to be a conflict here over a model of a universal disease classification system that is applicable in every kind of clinical setting. The needs for classification or diagnosis are very

different. The population served by the analyst in practice is a very different population from Pilgrim State Hospital. Whether one needs one comprehensive kind of system that will meet universal needs, or whether one takes the view that the diagnostic system should have some kind of practical utility for prediction or treatment is at issue. One can list criteria for psychoanalytic practice that are meaningful diagnostically but might be quite different from those needed in a large State hospital. This is brought out also in the history prior to the 1952 classification, when the older system only seemed to account for 10 percent of the population of people encountered in the military.

Dr. SANDIFER. I want to make one comment about multiple diagnoses. In a recent series of films we have tried this with our diagnosticians, presenting them with a list of broad categories and a five-point scale. They could make some statement about each category on the five-point scale, so that nothing was excluded. This was the same group who had previously seen films and were confined to making a single diagnosis. The clinicians' response to having opportunity for multiple diagnoses has been a very warm and positive one. It is their belief that it is a more valid description of patients than a single system.

¹ See opening discussion of paper by Dr. Temkin.

Epidemiology and Medical Care Statistics

Ernest M. Gruenberg, M.D., Dr. P.H.¹

Introduction

This paper was originally planned as a joint paper with Dr. Morton Kramer. Unfortunately his duties prevented Dr. Kramer from participating in its preparation so some important biostatistical and medical care concepts are poorly presented and may well be erroneous. Dr. Kramer will provide his views in another paper.

The two topics, "epidemiology" and "medical care statistics" have certain common needs in classification, but they are by nature different fields and so have certain different needs. This paper is therefore divided into three parts: "Part I: The Role of Classification in Epidemiological Research in Psychopathology"; "Part II: The Role of Classification in Medical Care Statistics Regarding Psychopathology"; and "Part III: The Common and Divergent Needs in These Two Fields."

An appendix on "Classifications and Nomenclatures of Psychiatric Conditions" offers some thoughts on a classification of classifications to which reference is made in all three parts.

Part I: The Role of Classification in Epidemiological Research in Psychopathology

The first part of this presentation, which is terse and formal, attempts to confront

the reader immediately with important general propositions and definitions in the field of epidemiology in general, which are no different from those in the field of mental disorder epidemiology. Later, the systematic outline is "fatted up" with more discussion.

1. Definition of the Field of Epidemiology

1.0 "Epidemiology" refers to the science which studies "the mass phenomena of disease" (Greenwood, 1935) by determining the distribution of conditions or diseases and the factors which determine these distributions (Lilienfeld, 1959); that is, it is "the study of the distribution and determinants of disease prevalence in man" (MacMahon, 1960). The analysis which epidemiology makes of these findings results in a "medical ecology" (Gordon, 1952). Epidemiology relates observed distributions of disorders to the environments in which people live—the physical, biological and social environments.

2. Units

2.0 The *prevalence rate* is by definition the number of cases present at time t ÷ the population of which the cases are a part (the population in the denominator includes the cases).

2.1 The prevailing cases at time t have arisen during various prior time periods. The rate at which cases of a disorder start is measured by

2.2 The *incidence rate*, which is defined as the number of cases starting during a

¹ Professor of Psychiatry, Department of Psychiatry, College of Physicians and Surgeons, Columbia University, New York, New York. Supported by Foundations' Fund for Research in Psychiatry.

specified time period ÷ the population at risk. This may be written:

$$[\text{Incidence rate}]_{t+1}^t D =$$

$$\frac{\text{Number of cases starting between } t \text{ and } t+1}{\text{Population at risk}}$$

2.3 *Duration of a case* is defined as the time between the onset and termination of the case.

2.4 If (a) Incidence Rate of a condition has been constant for a period of time (> than maximum duration of the condition) and (b) the average duration of cases has been constant during that period, then

$$\text{Prevalence Rate} \approx \text{Incidence Rate} \times \text{Average Duration}$$

where Incidence and Duration are measured in the same units (e.g. months, years)

2.4 *By definition.*—Prevalence Rate ≤ 100 percent. But if average duration of a condition is less than the time units in which an Incidence Rate is measured, or reported, then the Incidence Rate can exceed 100 percent.

For example, in some commands in World War II annual incidence of G.C. exceeded 100 percent in some time periods even though a minority of the command ever contracted G.C. Obviously, the URI annual Incidence Rate can easily exceed 100 percent.

2.5 *Other quantification devices.* — There are other ways of quantifying the amount of disease in a population, such as contingency risks, expectancies and period prevalences. But for the purposes of this discussion it is not necessary to elaborate on these other methods of quantifying the rate of occurrence of diseases in populations, since the basic issues for all can be specified with those already described above.

3. Differences in Rates of Occurrence

Epidemiology studies distributions of disease prevalence and their determinants. Hence only *differences* in rates of a condi-

tion's occurrence are of importance (there are situations which will occur to the thoughtful reader, later on in particular, in which these differences are measured as 0, and they can be of importance in drawing inferences too).

Formally, if P₁ and P₂ show R₁ and R₂ (that is, population 1 and population 2 show rate 1 and rate 2) for a condition, then if R₁ < R₂ there must be an explanation for the difference. This explanation will be sought in the differences between P₁ and P₂.

4. Agents, Hosts, Environments

It is commonly said that these differences are due to—

4.1 the nature of the members of P₁ and P₂ ("host" characteristics), or

4.2 an "agent" which is more prevalent in P₁ than P₂, or

4.3 the "environment" in which P₁ lives being more conducive to the development or prolongation of the condition than the "environment" in which P₂ lives.

5. Classifications Implicit in the Basic Terms of Epidemiology

5.1 *Cases of a condition.*—Here we are talking about the classification of psychopathology and the identification of individuals. This topic will be taken up in more detail in section 9.

5.2 *Populations.*—Obviously the classification of populations is necessary in carrying out the maneuvers outlined above. The usual methods of classifying population differences in epidemiology are discussed in paragraph 8.3.

5.3 *Times.*—The time interval during which an incidence rate is being counted or computed, the point in time at which a prevalence rate is computed, the time of onset, and the time of termination all require temporal classifications for epidemiological research.

5.4 *Hosts.*—As will emerge, there are many ways of classifying hosts; no standard

way of classifying them is indicated for epidemiological research. Some classifications of hosts are customary and have particular advantages but the selection of a classification scheme for classifying hosts depends on the question being asked and on the investigator's ingenuity and wisdom.

5.5 *Agents*.—The classification of agents emerges in part from epidemiological work and in part from epidemiologic hypotheses. Again, there is no standard way of classifying agents. However, it is common to classify agents as living and nonliving, and to classify nonliving agents as physical or chemical, and to classify deficiencies as agents. Those who enjoy scholasticism can have long debates on, "Are specific genes transmitted from parent to child, or arising as a mutation during oögenesis or spermatogenesis, agents and, if so, of which class?"

5.6 *Environments* are classified as physical, living and social. It has been said that while current difficulties in epidemiological research regarding mental disorders appear to be due to difficulty in classifying disorders, the classification of disordered states of mental functioning is far more advanced than classifications of the social environments in which people live and about which etiological hypotheses abound. This attempt to contrast the sociologists' and psychiatrists' capacities to classify the phenomena in which they are expert may not be justified. But it is a colorful way of dramatizing the fact that classifying environments is a prerequisite of epidemiological research. The needs of psychiatric epidemiology will not be met by classifying disorders alone. One must be able to classify the conditions under which these disorders arise or fail to arise.

6. The Uses of Epidemiology

Seven uses of epidemiology can be distinguished (Morris 1964). First, knowledge regarding historical trends helps to distinguish increasing disorders from disappearing disorders; second, community diagnosis

of the size, location and distribution of a condition aids in planning health programs for the community; third, by accumulating cases over a time period in the lives of individuals, individual risks can be calculated (a basic tool in calculating insurance premiums) and knowledge of contingency risks aids in estimating the effects of host factors in determining the distribution of cases; fourth, knowledge of the attributes of cases not in treatment enlarges the clinical picture by making our concept of a disorder less dependent on the clinician's limited perspective on cases; fifth, occasionally new syndromes are identified because clinically dissimilar cases are found to arise from a particular common background, or because clinically similar cases are found to arise in two or more distinct sets of circumstances; sixth, the "working of health services" can be studied in terms of their successes and failures, their selection of cases for treatment, and for their deleterious effects on the people they seek to serve; seventh, in the search for causes of disorders, data on the factors associated with the distribution of a disorder supplement laboratory and clinical data in the elucidation of causal mechanisms; at times the crucial breakthrough in our understanding of the way in which a condition is caused is made by epidemiological inquiry (this occurred with cholera, pellagra and lung cancer).

(The above lengthy sentence expounds a "classification of the uses" to which epidemiology is put. This classification, unlike the others mentioned in this paper only facilitates thinking and talking about the field of epidemiology and is not necessary for the conduct of epidemiologic investigations. It is, however, "useful".)

7. Special Classification Needs Arising From These Uses

7.1 *Historical trends*.—Long-term secular trends in disease incidence are difficult to appraise because this appraisal requires the diagnostic standards of one time and

place to remain the same for another time and place. It is difficult to distinguish improving diagnostic acumen from an increasing rate of occurrence of a condition (in some instances improving diagnostic acumen leads to the disappearance of a condition, such as typhoid-pneumonia, apoplexy, mania). Hence the need for stability in classification systems of disordered states becomes very apparent if one is interested in secular trends in the rate at which a given condition is occurring in a community.

Another special difficulty arises in defining the same population twice in two different time periods, when much has gone on with respect to the population in between. A classical example of this problem arises in a study of Goldhamer and Marshall regarding the first mental hospital admission rates for schizophrenic psychoses among residents in Massachusetts from 1843 to the 1940's. A great deal of methodological detail went into this investigation (which did not show any rise in first admission rates when appropriate correction factors were taken into account). However, the question remains as to what it means to speak of the population in Massachusetts as a sample studied over a century in time. Is it the same population? Is it a similar sample of a larger population at these two periods in time? It is not the intent of this article to answer this question but to raise the question to indicate a classification problem that arises in the course of studying secular trends regarding mental disorders.

7.2 Community diagnosis.—In measuring the size, location, and nature of a public health problem, epidemiological methods are used to count cases in a community. Epidemiology thus acts as the intelligence branch of public health practice, aiding in the planning of health services, preventive, curative, and rehabilitative by appraising the enemy to be attacked. To do a useful job such studies

must discriminate cases that could benefit from curative or rehabilitative services from other cases with the same diagnostic condition. Hence this practical problem of planning health services requires a classification of sick people by whether their condition is regarded as susceptible to the current technology of treatment or rehabilitation. Ordinarily neither morbidity surveys, counting numbers of sick cases, or mortality rates make this distinction. Likewise, tabulations regarding the incidence of conditions in a population do not in themselves discriminate between conditions which no one knows how to prevent from preventable conditions. The onset of a condition which is 100 percent preventable represents a failure of public health practice, but this is not true of the onset of conditions like arteriosclerosis, which cannot be prevented with present techniques. In the field of mental disorders the American Public Health Association's "Guide to the Control of Mental Disorders" undertakes to classify mental disorders in terms of their preventability, the presence or absence of effective treatment techniques, and the availability of rehabilitative techniques of known effectiveness.

7.3 Individual risks.—How much does smoking cigarettes increase the likelihood of lung cancer? Of coronary heart disease? How much does losing a mother in early childhood increase the risk of neurotic symptoms in childhood? Of neurosis in adult life? Of psychoses in adult life? How does the risk of bearing a mongol child rise with the age of the mother? With the age of the father? Such calculations of contingency risks according to special characteristics of the people involved handle time in a different way than incidence rates do, although calculations of contingency risks can always be looked upon as cumulative incidence rates over a period of years.

Calculations of individual risks highlight sharply the importance of being able to

classify the special characteristics or experiences of the population being studied. An important example today is the relationship between people who have never been married and people who have been married with respect to mental hospital admission rates and retention rates in hospitals. The characteristic "never married" becomes a key classification. It is obvious that this designation or classification of individuals must be applied by the same rules to those who have the hospital experience and those who do not in order to compute a contingency risk. This is often easier said than done because ordinarily information about the people in medical care and the people out of medical care is not collected at the same time and in the same way. If the Census Bureau asks of an institutionalized population what its marital status is at the time of the census, it then becomes possible to look into those relationships more exactly. Another practical classification which sounds simple but in fact is difficult is the classification of people according to the season of birth (which is linked to hypotheses regarding unfavorable fetal environment). The difficulty is knowing what the distribution of months of births is among people who are not cases and are members of the same population from which the cases were drawn. This is not so much a problem in classification as a problem in data gathering and population sampling. But the classification of experiences and characteristics of individuals underlies the analytic application of contingency risks measurements.

7.4. *Enlarging the clinical picture.*—Enlarging the clinical picture of a disorder in breadth is accomplished by locating other cases of the same disorder who are not in medical care but are members of the same population that the cases in medical care came from. This sometimes gives a broader perspective of the spectrum of the disorder than is otherwise found when only cases from hospital or outpatient

work are accumulated. It requires defining the class of cases in such a way that cases can be identified even though they are not in treatment. This, of course, is fairly simple in a series of steps following X-ray screening for tuberculosis. In such instances, the screen process ends up with the same criteria as the hospital diagnosis would. In order to get a wider picture it is sometimes necessary to modify the usual clinical diagnostic criteria so that they are applicable outside the clinical situation. If one then wishes to be precise, one can apply the same criteria to people who are in clinical care. For example, one can confine the diagnostic criteria for severe mental disorders in the aged to information obtainable in an interview by a trained interviewer. This makes it possible to compare cases who meet these criteria who are under medical care with those who meet these criteria and are not under medical care and to get a broad community picture. The point being emphasized here is that the need for this kind of information does change the need for classifying criteria in order to meet the practical problem of ascertaining cases in a consistent fashion among those in care and those out of care. (Enlarging the clinical picture in depth means finding out about cases who have the same condition without as many or as severe manifestations; one must undertake to identify the subclinical cases.) Such investigations are called for when there is a concept of a disorder which is only sometimes manifested by clear evidence of illness or disability. This is easily grasped in connection with a condition like poliomyelitis virus infection in which understanding the dissemination of the virus in populations requires identifying many individuals who have been infected with a virus but never been identified as poliomyelitis cases. This condition (poliomyelitis virus infection) is only rarely accompanied by a manifest disease, therefore, cases become classified according to

the presence of evidence of virus infection determined by the development of antibodies. Psychiatry has long had the concept of subclinical cases of such conditions as schizophrenia, manic-depressive psychoses, senile deterioration, mental deficiency. These conceptions derive largely from clinical experience in taking histories of cases in which the clinician recognizes stories regarding early manifestations unrecognized by the patient or his family and also in following up cases no longer in clinical care. Efforts to apply the concept of subclinical cases in psychiatry have not all been unsuccessful but great difficulty arises in this field because the definitions of subclinical cases have been inadequately worked out for this kind of work.

The concept of a subclinical case must be sharply differentiated from the concept of a remitted case. Manic-depressive psychoses illustrate this problem. A manic-depressive patient between episodes, who does not have manifestations of manic-depressive psychosis at that point in time, is not a subclinical case but is an individual who is not a case and is between episodes of what is conceived of as a psychotic disorder. This problem of classification is confounded if the concept "manic-depressive psychosis" includes the concept of a long-term condition with episodes of relapsing. This concept is somewhat analogous to the older situation with malaria in which an individual was not ordinarily ill except seasonally, although as knowledge advanced people were found to be harboring the malarial parasites even during their periods of remission. When such a concept is applied to a condition such as manic-depressive psychosis, there is a sharp difference between the system of classification so used and a system which identifies episodes of psychotic disorders of a depressive or manic kind to be studied in terms of the contingency risk of repeated episodes in the people who have already had one episode.

These comments are not meant to solve any problem respecting the classification of cases in epidemiological studies but are intended only to point to specific varieties of problems in disorder classifications which arise in epidemiological studies.

7.5 *Identifying new syndromes.*—It is clear from earlier examples that epidemiology often seeks association between prior experiences of individuals and the development of disorders. When this is done, individuals are classified according to their prior experiences and then according to whether or not they develop manifestations of the disorder being investigated. It is sometimes helpful to turn the procedure around and enumerate the manifestations of disorder exhibited by individuals who have had a common experience. Psychiatry is currently going through the process of digesting the information obtained when investigators began to ask what disorders occur at a higher frequency in children who have survived in unfavorable fetal environment during pregnancy. The work of Pasamanick, Lilienfeld, Knobloch and others, which has shown that such children are at higher risk of convulsive disorders, reading disabilities, impulsive behavior disorders, mental retardation, and a number of other clinical syndromes forces us to regroup children according to whether or not they have a manifestation of a "brain damage syndrome." This changed perspective arises from looking at antecedents rather than consequences of a disorder or a group of disorders, and it should not be surprising that preoccupation with antecedents will lead to a different grouping of cases than preoccupation with current manifestations or treatment needs.

In such instances, epidemiological inquiry leads to a change in classification systems: epidemiological knowledge modifies the classification of cases.

7.6 *Working of health services.*—It might be said that epidemiologists look on health services as a particular facet of the

environment in which people live. Health services might best be approached from the epidemiological point of view as differing from other environmental forces only in that they are the result of conscious, purposeful, socially organized attempts to modify the incidence and prevalence of diseased states. As technology improves, these health services become more and more relevant to the distribution of disease in the population. From this point of view, the cases under care in a health service can be looked at by themselves in terms of sorting out whose health is modified after contact with the health service in contrast to similar cases who are not in an improved condition after contact with the health services. Such a study obviously requires criteria of improvement and it should be clear that these criteria of improvement in classification of people according to whether or not their condition has improved is absolutely necessary for such inquiries. (It also should be obvious that this is quite different from classifying people according to whether or not they have a given disorder.) This requires a classification of the "goals of treatment." The American Heart Association has done outstanding work in this respect and a number of chronic conditions have been approached from that point of view. Clinical trials and other kinds of studies being discussed in this symposium also require the same classifications so this point will not be elaborated here. The health services can also be examined in terms of which cases of the disorder come to the health services and which do not. A study of this problem requires classification of sick individuals according to their reasons for seeking medical care. (See appendix "Presenting Syndromes.")

The classification of the deleterious effects of treatment sometimes poses new and special problems. When infectious jaundice is a complication of salvarsan treatment for syphilis no new classification of infectious jaundice is needed. But when

inappropriate care of psychotic individuals is thought to lead to an entity called institutional neurosis (Barton) the consequence is a new method of classifying cases, as has been used by Wing et al. in a study of British mental hospitals and has been used by the author in a study of the "Social Breakdown Syndrome." Here again, attention to common consequences of what is regarded as a pathogenic experience or situation leads to a new grouping of cases using a new set of manifestations as defining criteria.

7.7 *Search for causes.*—The search for causes is, of course, the most important activity in epidemiology and its main needs have already been stated in connection with other uses of epidemiology. The need for a classification of cases based upon observable characteristics, the need for consistency of diagnostic criteria during the course of a study, the need to identify the "not case" by the same criteria as "cases," the need to sort populations by their characteristics, prior experiences, or types of environment in which they live, and the need to classify individuals as to whether or not they belong to the "study population" are all obvious needs. Their necessity is easily enough grasped.

The practical problems in finding suitable ways of meeting these needs are not so obvious. For example, it is common to try to define a study population as the residents of a given community (town, township, county, city, health district, etc.). In practice such a classification turns out to be neither exhaustive nor exclusive of all mankind. There are numerous individuals who are both residents of the selected area and residents of another area: hence the classification into "study population" and "not study population" is not exclusive. There are also numerous individuals who are not a resident of the selected area and are also not residents of any other area; hence this classification is not exhaustive. Efforts to elaborate the definition of the class of "residents of the

study area" so as to make the definition both exclusive and exhaustive are complex and difficult to apply in practice.

8. Epidemiologic Methods

8.0 *Special needs arising from epidemiologic methods.*

8.1 *Classification of cases and noncases.*—For this it is necessary to have objective criteria clearly specified for this classification. It is necessary to specify the observers and how they will apply the criteria. It is necessary to have information as to how observers vary in applying these criteria in classifying people. Five types of variations are known to occur: (1) Observer to observer; (2) one type of subject to another type of subject (e.g., men and women); (3) condition to condition of observation (in the clinic or in the home, for example); (4) time to time (different times of day, different days of the week, etc.); and (5) secular trends (do the observers become more expert? Do they become more sloppy?)

It is necessary to specify the methods of ascertaining cases, that is to say, the procedures that the observers will go through in order to come in contact with persons thought to be at risk. This is particularly important when comparing the results of one investigation with another and forming judgments as to the implications of the results of several studies. Because of these special needs, at various times epidemiologists have tended to become expert diagnosticians on the conditions which they are investigating. This has been particularly true of the infectious diseases where the epidemiologist not infrequently acts as a consultant to the clinician. As has been pointed out, there are needs for specification in classification of the time of occurrence, the specification of onset and termination, the classification of cases according to the episodes. (The levels of severity must be classified, again by objective criteria with information regarding ob-

server variation and information regarding the ascertainment methods.)

The classification of population differences for descriptive and other purposes have been mentioned: Age, sex (some people still think there are only two and that this is an exhaustive, exclusive, classification), occupation, marital status, etc., are all in use. Classification of people according to the experiences they have been exposed to, places of residence, places of occurrence of the disorder, places of work, membership in groups, societies, segments of a society, etc., are all in use and illustrate special needs of the epidemiologist.

9. Classifications of Conditions

This whole conference focuses on this topic and the epidemiologist apparently has a use for every classification any of the conference members will advocate. Any condition of concern to the clinician can be of concern to the epidemiologist. His needs for a consistently applicable classification of people according to whether or not they have a condition have been spelled out above. But there is no scheme of classification of conditions which is of more value to the epidemiologist than another.

In this sense, epidemiology is not an independent field of inquiry, but interacts with clinical diagnosis and laboratory diagnosis. The need of the epidemiologist for a classification scheme regarding the conditions being inquired into is obvious. A few examples have been given where the epidemiologist calls attention to a significant difference in cases clinically grouped together. A more recent example of this has been the laboratory findings regarding the difference in the type of chromosomal abnormality commonly found in the mongols who are born to young mothers as compared to the mongols born to older mothers. Further discussion of the

merits and demerits of various forms of classification occurs in the appendix.

Part II: The Role of Classification In Medical Care Statistics

Some aspects of medical care statistics have been dealt with in the previous discussion of epidemiology. There are, however, some special needs of medical care statistics for classification which should be pointed out. The most important and obvious of these special needs is that in medical care statistics each patient is a unit regardless of how many conditions he may have, whereas in epidemiological inquiries the units are people with a given condition; thus in inquiries regarding several conditions the same person can be counted several times. While this can occur in some aspects of medical care statistics, the ordinary needs of statistics regarding medical care tend to put an emphasis on the classification of patients into exhaustive, exclusive categories for the overall appraisal of the operation of the medical care facilities. Certain administrative needs call for some elementary statistics which are hard to gather consistently and tend to become a matter of preoccupation. Thus the number of admissions in a year, the number of persons admitted in a year, the distribution of durations of hospital stay, the conditions for which patients have been admitted and their distribution, the causes of death occurring in the hospital, are all basic to medical care statistics. The tendency to demand a primary diagnosis for the purposes of tabulation forces a classification based on the principle, "One patient, one diagnosis," which is discussed at some length in the appendix.

The need to specify duration of hospitalization in a consistent manner has turned out to be fantastically complex in psychiatric work. Because of the administrative structures of governmental hospitals patients are carried "on the books" who are

nowhere near the hospital; in fact, and very often are receiving no clinical care at all. Likewise because of some peculiarities of the outpatient activities of State mental hospital staffs it is not unknown for former patients to be receiving clinical care while "not on the books." The definition of a release from a mental hospital has been something to which Dr. Kramer has devoted a great deal of attention and made great progress in getting the various states to standardize this concept. This is a special need arising in medical care statistics if they are to be of any value at all.

Another classification which has turned out in practice to be extremely difficult but conceptually very important is the distinction between first admission and readmission. The difficulty arises from the fact that a hospital system tends to regard a person as a first admission if that system has never had that person as a patient before. However, many patients have been in mental hospital facilities in other hospital systems. Some mental hospital systems treat a transfer of a patient from one hospital to another as a discharge from the first, and a readmission to the second. Hence in medical care statistics there is a clear need for classifying durations of hospitalization and prior hospitalization. Medical care statistics also encounter very complex problems in classifying medical care facilities. Broadly one can classify State mental hospitals, voluntary, proprietary mental hospitals, general hospital psychiatric units, organized outpatient clinics, and private outpatient psychiatry. While there are difficulties in applying such a classification, requiring a definition of each class, this always appears to be a soluble problem until society comes up with a new kind of entity like the local mental health center. While the classification of facilities is in principle soluble by arbitrary decisions regarding boundary phenomena, the classification of psychiatric contacts presents more difficulty. Oper-

ationally it is fairly simple to classify people according to whether or not they have been seen by a psychiatrist during a specified time interval, say a year. But their reason for seeing a psychiatrist might be for a screening examination for employment or admission to a university, or the Armed Forces. This is easy enough to deal with if one sets up a special class of screening contacts, but decisions must be made as to whether the judgment of the psychiatrists at the time of the screening contact should be recorded and classified. Furthermore some people see psychiatrists because they are regarded as a source of information about someone who is a patient or is a potential patient. This is a clear enough classification when the distinction between patient and informant is sharp, but all who have had experience know this is frequently not the case. In addition, some sick individuals, in particular children, are appraised and treatment recommendations are made by psychiatrists who never come in direct contact with the children but only with the informants, either parents or teachers. How does one then classify the relationship between the sick individual and the medical care facility?

The classification of periods in contact with the health services are even more complex for outpatient-types of resources than for inpatient resources. In practice, administrative reporting has tended to set an arbitrary rule that if a patient has not been seen for 6 months he is a terminated case and when that case is seen again he is reported as newly becoming a case again. The need for such a classification is obvious. Medical care statistics need the same kinds of information about the personal characteristics of the population being served as are needed by epidemiologists. They need to know age, sex, marital status and often wish to know many of the same things that an epidemiologist would want to know in order to make analyses of whom the health service is seeing,

where and when. There is a strong tendency for medical care statistics to begin to inquire into what kinds of conditions personnel time is invested in medical care facilities. This kind of inquiry obviously requires not only a classification of the conditions which the patients have but also classifications of staffs and classifications of staff work units.

Medical care statistics frequently are used to trace the way in which the same cases use different resources during the course of their disorders. For this it is important to classify patients as people and to be able to identify each member of the patient population uniquely. This leads to problems in what is conventionally called record linkage which can be regarded as a form of person classification. The basic problem appears to be to make sure that the same person is identified only once in the patient population and that all records pertaining to him are brought together in a single part of a file. This classification, according to who the patient is, is important in its complementary form: that is, it is important that two different patients are not linked as though they were a single patient.

Medical care administration is a rapidly growing field and is entering into many interesting and complex problems of cost accounting, medical audits, and other forms of operational research on the health services. This brief statement is, therefore, undoubtedly inadequate to indicate the whole range of medical care statistics as they could be conceived in a broader context: Medical care statistics in particular, where the classification of the condition for which the patient is being treated is made by clinicians working on organized medical care services. The classification of conditions must be adapted to the ideology of the clinicians working on the services. There must be a place in the classification for the clinician to put each case. This need works counter to the logical desire to have an

exclusive system of classification, since different schools of clinicians will want freedom in allocating their cases. Hence official classifications developed for such services tend to be in the nature of compromises between what the clinician wants to be able to do in describing his case with a diagnostic label, and what the medical care statistician needs, which is a consistent system of classification. In my opinion, it is useless to seek a standard classification of conditions for medical care statistics which will satisfy all reputable clinicians and will satisfy the logical needs for a systematic classification. This topic is discussed further in the Appendix.

Part III: Common and Special Needs

Epidemiology and medical case statistics have common needs to classify disorders, but these are rarely independent of the needs of clinicians and other investigators.

Both fields need to classify cases in other dimensions—age, sex, place of residence, social group, etc. Both fields need to classify people into “cases” and “not cases,” to classify episodes by the times when they begin and end.

Both need to be able to classify levels of severity and stages of improvement.

Both need to keep the classification of disorders distinct from the classification of patients, but medical care statistics can endure a classification of patients better than can epidemiology.

Medical care statistics has special needs for classifying types of facilities and treatments. It also is more preoccupied (currently) with record linkage (classifying records according to the person referred to) than is epidemiology.

Epidemiology is preoccupied more with classifying for case ascertainment methods, with applying consistent criteria for classifying all members of a population by the same methods into “examples” and “not examples” of a given disorder, and into

“members” and “nonmembers” of subpopulations.

Emphasis has been placed in this paper and in the appendix on the wide variety of classifications of psychopathology which are of potential value; the particular value and limitation of each type of classification needs to be recognized.

APPENDIX

“* * * if our linguistic outfit is treacherous, it nevertheless is indispensable—it is not always new words that are needed, but a means of controlling them as symbols, a means of discovering to what in the world on any occasion they are used to refer (27) * * *”

It is not possible here to discuss in detail every psychiatric nomenclature. A rough classification of classifications and nomenclatures is offered to provide a framework for evaluating existing trends and highlighting current issues regarding the naming of categories into which we of necessity must group patients. Although the discussion begins with the problems encountered in classifying cases, subsequently the crucial distinction between classifications of cases and classifications of disorders is emphasized.

There are in fact several sets of nomenclatures in psychiatry. The psychiatrist's roles vary, depending upon which function he is fulfilling for the patient, the family, the community, or the medical specialty of which he is a part. These varied roles of the psychiatrist are used in this paper to group the frames of reference which different sets of names serve. This method is used to help put differing nomenclatures in perspective and to provide a specific frame of reference in discussing each nomenclature.

We start with (a) the way in which people enter a relationship with a psychiatrist, move on to consider (b) the roles the psychiatrist assumes in relationship to the patient and referring family, or professionals, and then to (c) the roles of the psychiatrist as discoverer and student of psychopathology, and only towards the end of this paper assume (d) the role of the psychiatric nosologist who evaluates nosological systems invented by other psychiatrists.

The basic terminology of psychiatry is the language of human beings describing distress, uneasiness, anxiety, fear, or some other complaint. Patients can express themselves only in words they already know. The art of helping the patient to describe his experiences, feelings and thoughts, without suggesting to him what we suspect, is the art of the mental examination. It is referred to here as a re-

minder that the local, current vocabulary of patients is, in a sense, part of psychiatric terminology.

The terms used by psychiatrists to describe what their patients present to them is the special vocabulary of psychiatry. These terms are supposed to describe what the clinician observes and what he infers. Thus the use of words to describe a patient's problem, condition or disorder is a practical necessity for the psychiatrist in everyday clinical work.

HOW THE CLINICIANS FUNCTION

As Predictor and Explainer

Two ancient functions of the physician are to account for the patient's complaints and to predict what will happen next. Thus the young man who complains that he is masturbating several times a week and that he is sure this practice will drain away his strength, or cause him to go insane, can be given a number of explanations (17). Whether the physician says directly that the fear is unrealistic or that it is justifiable and that the complaint calls for extended treatment, the patient will hear some corroboration or denial of his own prognosis. But the terms the clinician uses to describe this complaint to himself and to his colleagues should not carry with them the implication of a particular cause or prognosis. Throughout its history psychiatry has tried to find language which permits description without ascribing causes or consequences and simultaneously to develop knowledge regarding causes and consequences. This knowledge becomes expressed in terms too, but unfortunately sometimes the accumulated knowledge simply changes the meaning of the terms used previously for description. Thus, the young man's fear can be called a phobia or a delusion and each will carry its own implications. The best way to avoid these implications in a description is to use the patient's own words. It is no more appropriate to say he complained of a "delusion" than it is to say that a patient in a medical clinic complained of "thrombocytopenic purpura."

Some terms used to explain a phenomenon suggest an understandable cause, others a mysterious cause. Bewitchment or a god's anger are just as understandable to those who understand such things as are conflicts, infections, and complexes to those who understand them. A suggestion of mystery surrounds terms like "a condition," "heredity," "degeneration," or in frankly circular jargon such as "onanism" which simply provides a mysterious word for what can be said in simple language but encourages the illusion of explanation.

The more mysterious the explanation the less clear the implied prognosis. But certain terms in

psychiatry carry with them benign or ominous implications: a static condition is less frightening than a progressive one; a benign cause associated with complete recovery without complications implies the best prognosis. The previous sentence is full of classifying terms which have come to have specific meanings in medicine and psychiatry. Some special terms have been developed to incorporate a class of disorders with both a particular clinical picture and an anticipated course; among these are dementia praecox, mental deficiency, senile dementia and acute confusional state.

As Therapist

Treatment is another goal of the physician and requires that patients be classified as suitable for psychotherapy, for electroconvulsive treatment, as requiring restraints, or constant supervision. This, too, is a way of classifying patients.

Some treatments are directed at specific causes, such as in general paresis, pellagic psychosis, and thyrotoxic psychoses. The attractiveness of classification by cause is partly due to an inclination to believe that treatments directed at known causes are both more effective and more scientific because more rational. This belief does not correspond with the known facts of medical successes and failures; tuberculosis, leprosy, and post-traumatic psychoses attest to failures of treatment long after causes were known. The effectiveness of quinine in malaria, of electroconvulsive treatment in depressions, and of the phenothiazines attest to the usefulness of empiric treatments in the absence of causal knowledge.

The successes which followed the identification of general paresis as a syphilitic infection have stimulated many to expect the same kind of sequence in other mental disorders. By analogy, making unconscious motives conscious has been expected to remove symptoms thought to be produced by repression. Some of the resulting theories have proven effective in successfully directing treatment and others not. Every new theory of disease provides a new basis for classifying disorders.

The paradigm of identifying a disease entity, finding the cause, and discovering an effective way of attacking this cause has had a powerful effect on psychiatric classifications. Attempts to identify common underlying processes have led to some outstanding successes, not only in elucidating brain syndromes such as Wernicke's encephalopathy and Korsakoff's syndrome, but also in elucidating psychogenic mechanisms which also group cases according to an underlying common feature.

The mechanistic implications of finding disease entities to treat rather than patients have driven some psychiatrists to develop classifications which

deny the reality of arbitrary divisions among such a group of underlying disease entities.

As Administrator

Mental hospitals have classified patients into disturbed, cooperative, assaultive, soiling, working, suicidal, and similar categories. These categories arose from the needs of the institution as seen from the administrator's position. They serve a definite purpose in spite of their weaknesses and dangers.

Psychiatrists outside the institution also act as administrators and always must classify patients in terms of their ability to live outside the hospital, their needs for supervision and their danger to themselves and others. He is the one who should recommend hospitalization when it is indicated.

The Discoverer

The physician who sees patients and classifies them according to symptoms, causes, prognosis, treatments, and administrative characteristics is always in danger—if he is alert—of discovering a previously unrecognized pattern. A group of cases strikes the clinician as similar in some respect no one had previously thought to be crucial. This leads to a new group—a new category—and to efforts to distinguish these cases from other cases.

A major advance in knowledge always requires the creation of new case groupings. When medical research made the advances that allowed identification of general paresis and mongolism it became possible to group cases having particular laboratory and physical findings. It was found subsequently that the mental picture in these conditions varied greatly.

Many revisers of psychiatric classifications have tried to group cases together which might have a common pathology. This was Morel's intention when he grouped together a form of *démence précoce*, Kraepelin's when he developed the group into dementia praecox and later Bleuler's (*Vide infra* "Syndromes of Common Cause"). Freud simultaneously was regrouping cases in terms of levels of fixation and predominant mental mechanisms used in symptom formation.

These are illustrative of the theoretical implications of new classifications and of the classifying implications of new knowledge.

But each new classification tends, for various reasons, to be formulated too categorically and too comprehensively. Classification can help bring order out of chaos. However, classifiers tend to make their classifications cover too much ground and to make each label carry an entire picture of causes and consequences. In reaction against the misunderstandings created by classifications, innovations arise which sometimes change the labels, sometimes introduce a new basis for classification,

and sometimes reject classifications of cases and advocate a unitary theory. The best recent exposition of this unitary theory of mental disorders is provided by Karl Menninger in "The Vital Balance" (22). The viewpoint expressed by Menninger and his colleagues recognizes all the syndromes previously recognized but arrays them in levels of decompensation rather than classes of different disorders.

NOMENCLATURES ASSOCIATED WITH HOW A PERSON BECOMES A PATIENT

As the physician's various functions lead to different classifications of the problems brought to him so the ways in which people come to the psychiatrist also produce a useful classification (28).

As an Applicant for Help

A person who seeks a psychiatrist's help has come to the conclusion that his own personality functioning is awry in some way. Seven categories can be distinguished: "Subjective sensations" that the patient feels abnormal, such as depression or elation, loss of control of thoughts, depersonalization, phobias, anxiety, obsessions, and illusions (such as *déjà-vu*); "patterns of behavior" that he believes to be abnormal—usually aggressive, suicidal, sexual, delinquent, or cowardly ones; "bodily complaints," such as headache, skin rashes, and palpitations, that may be seen by the patient as having an emotional cause or presented as physical in origin (when this occurs the patient's awareness of the emotional concomitants of his complaint is shown by his lack of surprise when a psychiatric history is taken); "impossible life situations" and conflicts such as marital disharmony; "failures to achieve what the individual feels he should" in school, courtship, work or personal relations, and lastly, complaints of "confusion" or "disorganization" where the patient may feel something is wrong and not be able to say what it is (he may begin his explanation in the distant past or refer to what someone else has been doing).

As Cause of Concern or Fear in Others

A member of the family, a friend or a public servant, such as a welfare worker, may bring the patient to the psychiatrist, saying that the patient is a danger to himself or to others or is unable to care for himself. This is a longstanding convention (9) and from it are derived the phrases used in many laws relating to involuntary hospitalization. In fact many complaints fall into these categories. Danger to self includes failure to take usual safety precautions, impulsive recklessness, and failure to carry out medical advice. Danger to others includes assaultive behavior or threats but

sometimes has to do with improper disposition of financial assets which the complainant sometimes regards as simply a danger to the patient, sometimes as evidence the patient cannot take care of himself, and, if he is an heir, sometimes as a danger to the complainant.

This classification—(1) danger to self, (2) danger to others, (3) unable to care for self—is not a scientific classification; because each complaint cannot be expected to fall into one of the three categories and into only one. This semilegal grouping is not likely to be taken to have a scientific character. Yet many classifications used in psychiatry are no more exhaustive or mutually exclusive and hence are not, strictly speaking, classifications of phenomena.

As a Cause of Trouble

People not thought to be a danger and who have not committed statutory offenses may be brought to clinicians for help because they cause trouble through excessive belligerence, making spectacles of themselves and embarrassing family or friends, irritating others or appearing excessively unstable. Such ways of functioning are sometimes seen as manifestations of mental malfunctioning and grounds for consulting a psychiatrist. These are disturbing ways of functioning which should not be confused with failures to function.

As a Lawbreaker

A person arrested for an offense may be brought to clinical attention because law enforcement officers, a judge or the defense attorney suspect he falls in a category of disorder which makes him ineligible for trial as unable to conduct his own defense or ineligible for conviction because of mental disorder.

This classification cuts across all other classifications in psychiatry. Sometimes if psychotic illness or mental deficiency can be found, psychiatrists feel secure in claiming freedom from culpability and, if not, they feel free to say the patient is sane. But society's decision not to hold mentally ill people culpable of offenses is something other than the classification of cases into mentally defective, psychotic, neurotic, and character disorders. The concepts involved here are more closely related to the specific act and the specific mental dysfunctions than the diagnostic classification. The rules for classifying people as capable of defending themselves or as culpable are not sharply drawn. They reflect social attitudes and concepts of mental disorders and are defined in the interactions between psychiatrists and judges. The M'Naghten rules depend on a concept of disturbance in the defendant's capabilities while the recent Durham decision introduced the concept that certain acts can be attributed to a mental disorder in some cases (21, 10). Neither has gained universal ac-

ceptance and one can predict that proponents of each method of adjudication will continue to argue (16).

As a Functional Failure

The complaint often states a failure to perform a particular expected function. The failure may be in bodily self-care, in household duties, in study or work, in recreation, in personal relations, or in sexual functioning. This list is itself a rough classification of such complaints. In an extreme form the person may stay in bed, be mute, and not eat. In minor forms he may be adjudged to be not trying to get into college although a top intellect, or to be avoiding the effort to make friends.

These complaints all reflect social standards which vary both in differing social groups as well as from time to time. Failures to function fall into two broad classes: loss of functions previously carried out and failure to take on those new functions assumed to be appropriate for each age.

The Accepted Bases for Classifying Cases

Patients have been classified by psychiatrists on the basis of a number of different characteristics. Menninger and his associates have provided a good recent annotated historical bibliography of psychiatric classifications (21). It is a mistake to consider these classifications as exclusive, exhaustive classification of the cases of concern to psychiatrists. A classification of members of a class of individuals (here psychiatric cases) is said to be "exclusive" if each individual can be assigned to only one class. It is "exhaustive" if every individual is classifiable into some class. We may call it "logical" in contrast to "arbitrary" or "conventional" if the rules for class assignment are based on a general concept of what is important and the rules follow a pattern of reasoning about the classes. Linneus' classification of living organisms is the paradigm of logical, exclusive, exhaustive classifications. The logic of this classification led directly to evolutionary concepts and thus was highly productive. (This type of classification was dubbed "colligative" by Whewell (31) to distinguish it from arbitrary groupings which are not presumed to reveal inner connections in the things grouped together.)

In medicine a practical classification of cases cannot be based on such principles because some cases are characterized by the locus of a lesion, some by an outstanding manifestation, some by an excessive or defective functioning pattern (hyper- and hypothyroidism, hypertension, phenylketonuria, fevers, etc.) and some by a known or assumed agent (tuberculosis, smallpox, rubella, St. Louis encephalitis virus, etc.). When a case can be classified according to its relationship to a known agent this is usually given priority in assignment to a class. While organ pathology is usually given

priority over functional abnormality, all classifications do not hold true to this position.

Outstanding Manifestations

A single outstanding manifestation observable in a group of cases has been repeatedly used as the basis for a class of cases. Obvious examples which have endured over a thousand years include epilepsy and somnambulism. Epilepsy continues to be an acceptable diagnosis, a focus for treatment, a subject for investigation and an acceptable complaint. By contrast somnambulism which remains an acceptable term to describe a manifestation of disorder and an acceptable subject for investigation, is no longer a commonly used diagnosis, but is thought of as always symptomatic of something else—and thus “only a symptom”—nor is it usually regarded as a proper focus of treatment. While epilepsy also is “only a symptom” it yet survives as a diagnosis, partly for practical reasons and partly as a matter of fashionable ideology. If we wished to insist on a firm attribution of somnambulism to some other condition with objective evidence of the kind needed to attribute epileptic convulsions to brain tumor or infection, we would have many cases of idiopathic somnambulism and might even have a subject for investigation—somnambulistic disorders. The fact that we have discarded the use of somnambulism in classification is not a matter of abstract logic but of overall judgment by the profession that it would be meaningless. Somnambulism is assumed to be evidence of at least one other disorder in our classification and even though the evidence for another disorder is very weak and its connection with the patient's somnambulism undemonstrated, we usually put the case in some other class. (But both the American Psychiatric Association and World Health Organization classifications permit classification by this symptom alone if there is no evidence pointing to another diagnosis.)

Another ancient category of classification is lycanthropy. This term has disappeared from the APA classification although its manifestations continue to occur clinically and it is in the “Index to the World Health Organization Classification.” The term may still be found. Because this is an unfamiliar term and the syndrome is not referred to in modern texts it is sometimes not even recognized or described in the clinical case notes.

Grouping cases together because of a single outstanding manifestation is sensible if it is part of an effort to classify the mechanisms by which that manifestation is produced. For example, Freud wrote a paper on jealousy, paranoia and homosexuality (13), another on a case of paranoia (12), one on a phobia (11) and so forth. Being preoccupied with the pathogenesis of symptom formation, he repeatedly selected an outstanding symp-

tom for investigation. Similarly, grouping cases together with an outstanding common manifestation always is reasonable if a particular form of treatment appears to have an effect on this manifestation; epilepsy is an outstanding example of this principle. It was the object of the first systematic search for a treatment by rational methods (23).

The objection to a classification of this kind is that it pulls together cases arising from many different causes. As knowledge advances, this deficiency becomes more and more obvious. In Tuke's dictionary (29) many symptomatic disorders are labeled and described; the extensive terminology used to describe special symptoms at the end of the 19th century is clearly recorded. Hinsie and Shatzky's dictionary is more modern (18). More sophisticated classifications have replaced that fashion. But classes characterized by a single outstanding phenomenon continue to be part of official classifications. Even today they are included in the American Psychiatric Association (1) and World Health Organization (32) classifications in such terms as convulsive brain disorders and the special symptom reactions.

Syndromes and Symptom Complexes

The recognition that groups of symptoms and behavior patterns frequently occur together leads to the creation of classes which reflect these groupings. This procedure is a step away from classifying by an individual outstanding manifestation.

Thus there are patients who are anxious and there are patients who are depressed, and there is a group who are both anxious and depressed; anxious depression is a frequent symptom complex which is distinguished from anxiety states and from other depressions.

A group of patients is sometimes characterized by an outstanding manifestation and then separated into subgroups on the basis of a secondary characteristic. Thus the mentally retarded have been repeatedly divided into those with and those without physical defects. This has proved to be a fruitful way of looking at these cases (15, 7).

The History and Course of a Case

Very often the acuteness, chronicity, progressiveness or intermittence of a manifestation is used to distinguish one group of patients from another. Thus recurrent depressions are separated from an isolated depression, chronic anxiety is distinguished from an acute anxiety state, progressive loss of vision is distinguished from sudden loss of vision and so forth.

Such groupings of cases remain at a relatively simple level and are closely related to grouping by complaints or what is regarded by clinicians to be especially important manifestations of disorder.

More complex characterizations of cases are spoken of as syndromes which take into account complaints, outstanding manifestations, course and situation in which the picture of disorder emerges. Syndromes are patterns of disorder which are observed to hang together in a number of cases. Examples in other branches of medicine are shock, cardiac decompensation, marasmus, acute abdomen, and fainting. Examples in psychiatry are catatonic stupors, fugue states, depersonalization, mental retardation, confusion, panic and paranoid states. None can be characterized by a single symptom complex but each is recognized by a common feature considered to be of importance in one group of cases in contrast to all other cases. The important common feature may be related to a common problem of management, an urgent need for a differential diagnosis, a suspicion that the constellation of disordered functioning is due to a specific cause, or simply because the syndrome presents a startling clinical feature to observers. Depression and panic long have been distinguished by psychiatrists; while only some psychiatrists have thought that all depressions have a common causal element (20). Only in recent years have treatments directed particularly at these syndromes become available. Classifying cases in this way is important in clinical work. Named syndromes are used in discussions among ward personnel to communicate the problems presented by cases; they summarize complex clinical pictures in single terms which serve as symbols facilitating thinking about patients. Often syndrome classification serves as the basis of new ideas concerning causal mechanisms and of new methods of treatment.

Presenting Syndromes

In a recent review of our current state of knowledge regarding effective means of modifying mental disorders a subcommittee, following the advice of Dr. Heinz Lehmann, found that the best way to specify effective psychiatric treatments was in terms of presenting syndromes responsive to treatment (2). This group listed as presenting syndromes usually susceptible of early resolution by appropriate treatment: Depressive, severe anxiety (panic), chronic tension state, paranoid, manic, and confusional. The fact that each is seen as calling for somewhat different treatments strengthens the view that patients showing each syndrome share something in common which differentiates them from patients exhibiting other syndromes.

Grouping cases by presenting syndrome when each syndrome is related to a common technology of treatment or prevention is obviously an advance over groupings which are purely descriptive.

Grouping cases by syndromes which have a common underlying brain lesion or psychological defect represents a still greater understanding of the disorders. When such a syndrome is first suggested it is an hypothesis in the sense that its author has noticed a group of cases he believes to have a common underlying disorder. Earlier the outstanding examples of Wernicke's encephalopathy and mongolism have been mentioned. In both, the hypothesis has been confirmed by identifying a common underlying disorder. Wernicke himself described the lesions associated with the syndrome bearing his name. Langdon-Downe was dead many decades before the trisomy underlying mongolism was demonstrated. In the interval mongolism represented a class of cases which had many characteristics in common but also had many variable manifestations. It remained a category of uncertain boundaries until associated with the finding of trisomy-21.

Syndromes of Common Causes

A further step in grouping cases occurs when an attempt is made to group cases because their disorders are due to a common cause. Initially such a grouping is also hypothetical and may or may not lead to the identification of a common cause. In psychiatry the group of cases classified as general paresis has been the most successful such hypothesis because, in fact, it became associated with central nervous system syphilis. This victory, however, required a regrouping of cases because some patients who appeared to have the syndrome of general paresis did not, in fact, turn out to be infected by spirochetes and other patients who had only slight suggestions of general paresis as a syndrome did turn out to be experiencing a central nervous system infection with syphilis. Hence identification of an underlying lesion or common cause by laboratory techniques leads to a regrouping of cases in which the clinical picture is more varied but the boundaries more precise. In practice, investigations are needed to determine what syndromes are produced by the common cause and how distinctive they are. The relevant investigations regarding trisomy-21 are still going on.

Failure to distinguish between syndromes known to reflect a common underlying condition or cause from syndromes believed to reflect such a common cause leads to confusion. Efforts to crystallize such syndromes during the last century followed Morel's observation of a group of young cases with deterioration of mental functioning which he called *démence précoce*. Kraepelin built this syndrome into an "entity" characterized by its deteriorating course. E. Bleuler, in 1911, in perhaps the most brilliant exercise of this type, found that cases

which fit Kraepelin's description of dementia praecox not infrequently showed remarkable restoration to their previous state of functioning and that the dementing course was not associated with any particular group of presenting pictures. Using the insights of Freud and of other contemporary psychological investigators and theoreticians he divided the manifestations into a small primary group which could be attributed to the disorder directly and secondary manifestations resulting from the individual's efforts to compensate for them. As a clinical investigation attempting to find a way of grouping cases presumably having a common underlying disorder these efforts of Kraepelin and Bleuler have had an enormous effect. The "Group of Schizophrenias" have, however, tended to become transformed from a hypothesis into a diagnosis, used as though it was known to reflect some common cause. Common usage of a term like "schizophrenia" as though it represented a disease with a single cause does not replace evidence and such usage—by providing a feeling of assurance—conceals our ignorance and prevents us from seeking answers to questions which are not clearly asked.

The thought processes which accompany a statement such as "this patient has dementia praecox" or "this patient has schizophrenia" are sometimes accompanied by a failure to relate to the person who is in distress and to deal with the specific problems he presents to the clinician.

Partly because of this damaging effect of such classifications, Adolf Meyer introduced an entirely new way of grouping cases and syndromes. He was one of many scientists and philosophers who were deeply impressed by the observation that matter is organized into a hierarchy of levels of complexity and that each of the emerging fields of science tended to specialize in one or another of these levels. Thus atoms are grouped into molecules which have properties quite unlike those of the constituent atoms. Molecules are grouped into liquids, solids, and gases that have properties quite unlike the properties of individual molecules. The transformations of atom groups to produce other molecules are studied by the chemist. The formation of living cells by special organizations of molecules produces matter with properties unlike the inanimate aggregations studied by the physicists and chemists. Groupings of cells into certain organized patterns produce organisms with properties different from single cells. The nervous system in higher animals leads to classes of behavior not seen in organisms without nerve cells. Man has a nervous system which itself is organized into hierarchical strata of functional connections, which lead to mental life with symbolization and consciousness and self-consciousness when human

organisms live in social groups. Meyer, deeply impressed by how much was still unknown and by how much of mental disorder could be understood in terms of mental life, developed a classification based on the level at which the disorder had its origin. He grouped together conditions which had an origin below the mental level of functioning and which interfered with the capacity of the central nervous system to provide an adequate basis for mental life. There remained many syndromes which he believed would have to be understood in terms of disorders in mental development and sometimes in terms of the social situation in which a person grew up or later found himself. These he grouped according to the extent of disordered manifestations in the mental life of the patient. Thus those who showed no prior manifestations of disorder in mood or in their thinking but whose disorders were confined to a complaint or a specific problem were grouped into "part disorders" or "complaint disorders" under the term "kakergasia." Those whose ideas of the nature of reality were seen to be disordered were grouped into sweeping, disorganizing disorders for which he coined the term "holergasia." This overall classification divides cases on three general principles: first, whether the disorder arises at the central nervous system level or in personality functioning; second, in the latter group as to whether the disorder affects the personality as a whole or is confined to a single symptom, and, third, in disorders arising from malfunctioning of the central nervous system whether congenital and, if not, whether acute or chronic (26).

While this grouping of "reaction sets" does not differ greatly from the nosology of Kraepelin, the implications were seen as radically different.

Since disorders of the central nervous system are associated with disruptions of functioning called sensorial such a classification of cases is closely linked to the parts of the mental status examination which Meyer introduced into American psychiatry in a systematic way. These parts are: Appearance and behavior; complaint; present illness; mood; ideation, trends (content); grasp (sensorium), "support disorders" (of the central nervous system) produce mental dysfunctions revealed as sensorial defects in the mental status. Other cases, without sensorial defects, manifest sweeping disorders of content and make up another large group. Those cases with mood disorder but without sensorial or content dysfunction (except those accountable to the disordered mood) form a third group. Those with no marked sensorial defect, content manifestations, or mood disorder but with symptomatic complaints form a fourth group (the symptomatic neuroses). A fifth group of cases are characterized by their behavior rather

than by sensorial defects, content disorders, mood disorder, or symptomatic complaints.

This classification has the virtues of (1) distinguishing cases believed to have a disorder arising from central nervous system malfunctioning or structural change from other cases and (2) providing fairly simple rules for classifying other cases in terms of clinical manifestations observable in a psychiatric interview. Furthermore it does not require a knowledge of causes and is adaptable to advancing knowledge. Its greatest weaknesses are (1) that it has no way of incorporating clinical pictures which change over time and (2) it assigns priorities in classifying cases in a fairly arbitrary way that emphasizes thought content before affect or mood, mood ahead of symptomatic complaints, and behavior last. In view of these weaknesses, syndromes classified in this manner have been referred to as "reaction sets" to avoid the implication that the classification designates structural or causal factors.

"Official Classifications"

In this form it became the foundation for the U.S. Army classification (30) and for the American Psychiatric Association's official classification. It does not greatly differ from the World Health Organization's classification. The most important difference between the American Psychiatric Association's classification and the World Health Organization's classification—in the opinion of this author—is in the instructions regarding priorities in assigning a case. The American Psychiatric Association emphasizes an assignment to a "brain syndrome" category if there is any evidence of "impairment of brain tissue function" (1). The World Health Organization gives no explicit instructions; in practice, this probably emphasizes assignment according to the manifestation of predominant importance in bringing the case to clinical attention (32).

A conspicuous consequence of this difference is the tendency in Europe to classify some elderly persons as depressions who in the United States are classified as suffering a chronic brain syndrome.

The efforts to make official classifications comparable are important since they bring to light varying usages in different countries of similar terms. Unless the professional world knows which terms have the same meanings in different countries the scientific literature of each country can be used only by people in that country. The Health Section of the League of Nations played an important role in standardizing key aspects of disease nomenclature and this is now an ongoing function of the World Health Organization. The American Psychiatric Association is one of the few important national professional bodies which has not as yet

accepted the World Health Organization terminology. Current efforts to produce a revision of the "World Health Organization Classification" which will be accepted by all member nations (including the USA) are promising. But an international agreement to confine patient assignment to an agreed upon list of names does not give assurance that these names will be used in the same way nor does it imply that the basis of classification is everywhere the same. An agreed upon list of names represents only a step toward a standard nomenclature and a standard classification. Definitions of terms, instructions for assigning priority, and explicit statements regarding the basis for assigning cases to categories are additional tasks.

TWO CURRENT SYNDROMES

Advances in understanding are often signalled by the introduction of a new term to describe a group of patients who come to be seen from a new point of view. Useless invention of new terms clutters up professional literature with words which add nothing to our understanding and interferes with clear thinking. On the other hand, a partially recognized phenomenon often can only be clearly described by the creation of a new term, such as "institutional neurosis," "social breakdown syndrome," or be brought to prominent attention by the invention of a dramatic term such as "Munchausen Syndrome."

New names sometimes cause distress among those who wish for a stable world. But as our understanding grows, new insights frequently develop which demand new designations. (Clearly, making new names is not the same thing as developing new insights.)

Munchausen Syndrome (3, 4): This peculiarly named troublesome condition has received more attention since its clever designation than it had previously. Polysurgery was not unknown but frequently remained unrecognized as a mental disorder. The greater breadth of Munchausen Syndrome calls to attention the special function hospitals and doctoring play in the mental life of some people. The capacity for simulating the need for surgery as well as other medical procedures combined with the determination to be the object of these procedures has become a more and more conspicuous phenomenon. In the absence of a suitable term a new syndrome was designated by this title. Its designation facilitated communication between internists, surgeons and psychiatrists and made it easier to instruct students and nurses. It is a clinical syndrome which should be recognized and a name was needed to designate its occurrence.

Social Breakdown Syndrome

Chronic withdrawal, helplessness, soiling, or belligerence have come to be seen by many psy-

chiatrists as secondary consequences of institutionalization. The term "institutional neurosis" (5, 6) was invented to cover such institutionally produced secondary disorders. Some observers have thought these manifestations to be secondary to other social processes than institutionalization and therefore believe the designation by the epithet "institutional" as too limiting. Hence the term "social breakdown syndrome" was used (2) to describe a group of patients with deterioration of personal and social functioning supposed to be due to the way in which they and their primary disorders were managed in their medical and social environments. This syndrome is characterized as "sociogenic" and hence expresses a hypothesis, or group of hypotheses, which need testing.

While much current opinion supports this hypothesis or is consistent with it, it will remain a hypothesis until better data are available. Its naming and specification make such investigation possible.

These two syndromes of current interest are mentioned to draw attention to the fact that new names for syndromes play a role in psychiatry and to make the reader aware of the fact that the identification of new syndromes is a necessary step in reorganizing our knowledge of mental disorders.

SOME SPECIAL ISSUES

Cases or Disorders?

The previous discussion concentrated on the classification of cases. This exercise demands a class for each case and that each case fall in only one class.

"One thing is certain. We have to get away from the idea of 'one person, one disease.' Where would general pathology stand if it had to conform to the reports of the so-called cause of death, without qualification?" (24)

If a classification seeks to classify disorders this constriction is foolishly severe. If the classification tries to group the ways in which mental life can be disordered there is no a priori basis for assuming that each person is disordered in only one way. Nor is there any reason to assume that all possible disorders have previously been classified and to deny that some cases may be disordered in a way never previously observed. During the rapid advances in knowledge of medicine, surgery, and pediatrics, the strong tendency in leading medical centers to insist that all of the signs and symptoms of a patient be accounted for in a single diagnosis led to progress and the identification of new pathological processes. But many of the single diagnoses were acute and fatal, and as they have yielded to advancing technology, more and more patients are seen with multiple chronic diseases. In psychiatric services, it is

obvious that some people have both chronic brain syndromes and severe depressions, neither presumably the consequence of the other. In our mental hospitals patients who have had long-term schizophrenic syndromes can be observed to develop arteriosclerotic brain syndromes in addition.

These illustrations point up the difference between a classification of cases and a classification of patterns of disorder. Increased understanding is expressed in improved classifications of disorders. Easy administration and easy statistical tabulations arise from systematic classifications of patients which permit of only one class for each patient (33).

Definitions: Operational and Deictic

A class is designated by a name, by the definition of the class, and by the rules for including a phenomenon in the class. So much progress in so many fields has been associated with operational definitions that such definitions have achieved exalted status. Operational definitions avoid the problem of understanding the nature of the subject of discussion and make it possible to deal quantitatively with phenomena otherwise difficult to handle precisely. Thus the idea of force is not easy to define but, by defining a dyne as the force necessary to produce an acceleration of one centimeter per second per second in a mass of one gram, wonders were produced in physics. To define a case of mental disorder as a person of clinical interest to a psychiatrist is not comparable. It makes it possible to count cases by asking psychiatrists to name those who hold their clinical interest. But it adds nothing to the concept of mental disorder and obscures the fact that some people with a mental disorder are unknown to any psychiatrist. Operational definitions, then, are not always helpful—some operational definitions are circular, imprison our thinking, and obscure issues.

The simplest way of defining a disordered way of functioning is to show a group of cases. A group of cases, each of whom describes a thought or idea which he says keeps troubling him in a meaningless way, can characterize the disorder: we can then say that we mean by "obsession" the disorder which all of these people have. Many clinical phenomena are defined in just this way in every branch of medicine. Such definitions are referred to as "ostensive" or "deictic." The definition of concepts and phenomena by illustrative examples only is very common in psychiatry.

Deictic or ostensive definitions produce ad hoc classes which cannot be expected to correspond to logically separated categories. Logically separated categories need verbal specifications, but all such specifications depend on assumed terms and concepts, and these are often characterized by illustrative examples or loose verbal descriptions.

Hence every system of nomenclature depends upon some deictic definitions. Operational definitions, when appropriate, reflect an advanced level of familiarity with the phenomena, and when properly selected can lead to great strides in analyzing relationships.

Nosology, Classifications, and Nomenclatures

The general theory of classification is called taxonomy, but biologists use "taxonomy" to refer to the classification of living organisms. (Gregg refers to this as "taxonomy proper" which he contrasts with "methodological taxonomy" which is the theory of taxonomy—the classification of living organisms (14).)

Nosology is the classification of diseases and tends to be used in this way even when the classification includes syndromes not thought to be diseases. A nosological scheme classifies disorders, i.e., a variety of natural phenomena, and not cases. For many clinical services statistical classification of cases is necessary. The earlier discussion in this appendix was designed to make it clear that the classification of cases, even when based on a nosology of disorders, cannot be the same as a classification of disorders.

The nomenclature of a profession or field of work is the body of names used to refer to things, events, or actions by members of that profession. The word is also sometimes used to apply to the names of phenomena when classified in a particular way. By this usage nomenclature tends to become synonymous with classification, and, when applied to mental disorders, to nosology.

How Many Categories?

Classifications which become used as standard nosologies represent compromises between currently held views and currently emphasized phenomena. If a classification permits each viewpoint to classify disorders in its own way the classification must contain multiple classes for the same disorder.

In time, changing viewpoints, or newly acquired knowledge, lead to the creation of more categories. This strengthens the opposition to nosology by providing the argument that it contains too many categories and should be made simpler.

When a nosology is made shorter by grouping categories together it begins to lose meaning. The desire to make the complexities of life simple by reducing the number of recognized categories reflects a universal human desire for simplicity. "Nosology is preeminently one of the phases of medical pondering appealing to the sense of simplicity and order and striving to rise above the complexity of actual experience" (25).

There can be no final answer to the question, "How many categories?" Each answer depends on

how much of the complexity of psychopathology the makers of a classification wish it to reflect.

Circular Definitions

"To bring in prognosis as a leading feature in a nosological system is about as wise as to bring the issue of religious denomination into an election" (24).

When a disorder is described it is, of course, important to state what is known about the circumstances under which it occurs: the characteristics of people to whom it occurs (age, sex, family history, occupation, etc.); the usual course of events after the diagnosis is made and the available treatments and their effectiveness. From the viewpoint of nosology such factors should not have a decisive influence in differential diagnosis of individual cases. The definition of the disorder should not include such factors unless it is actually intended to create a category which separates examples of the clinical picture in men or women or with family histories from those without family histories and so forth. If such factors enter into the criteria for including a case in the category, research on the epidemiology of the disorder, or the natural course or the effect of treatment becomes meaningless. Thus if general paresis is defined as always fatal no case can keep the diagnosis who doesn't die; if the family history enters into the diagnosis of manic depressive psychoses such cases cannot be used to find the frequency of cases in family members as compared to the frequency of cases in family members of a control group; if combat fatigue is defined as a syndrome occurring only under battle conditions its presence in other circumstances cannot be studied. Manic depressive psychoses, reactive depressions, dementia praecox and mental retardation are examples of categories whose definitions contain outcome phenomena. Consequently, little information on the outcome of these syndromes is available, and the gathering of relevant data requires a restatement of diagnostic criteria.

Postpartum psychoses and menstrual tension are examples of syndromes defined as occurring in only one sex. (So was hysteria originally.) We may believe our knowledge or administrative needs justify behavior disturbances in children and transient situational disturbances in adults, but we should recognize that such categories close the door to certain questions.

The Dangers of Names

Coining a name—like coining a coin—may create something genuine or something with merely the appearance of value. Nouns appeared to Aristotle to have no element of assertion regarding the nature of the world except when linked to verbs. That this is an erroneous view of the effect of

names has become very apparent during the last hundred years (27, 19).

E. Bleuler listed "degeneration" as an example of autistic undisciplined thinking in medicine and psychiatry (8).

"A loose and insincere use of language leads not only to intellectual confusion but to the shirking of vital issues or the acceptance of specious formulae. Words were never a more common means than they are today of concealing ignorance and persuading even ourselves that we possess opinions when we are merely vibrating with verbal reverberations" (27)

A term's validity comes from its capacity to put phenomena together which have something in common, without suggesting any greater understanding than we actually have. Epilepsy, Munchausen's Syndrome, social breakdown syndrome, pseudoneurotic schizophrenia, and schizophrenia too, are each of value only to the extent that they do this.

Summary

Efforts to standardize nomenclature can be useful if they set realistic goals and if it is recognized that such efforts will never end so long as our understanding is growing. Standardizing nosological systems is attractive to administrative minds and is highly desirable when such systems reflect established knowledge; but if they seek to include all disorders they must, at any one point in time, be based in part on speculations and compromises; thus instead of making more conspicuous what remains unknown they tend to bury the unknown in neatly arrayed lists of labels.

Psychiatry cannot function without classifying disorders into named groups—such groupings and names are tools for communication, for thought and for research. As research yields new knowledge new meanings for old terms develop and new groupings of disorders require new names.

REFERENCES

- (1) American Psychiatric Association, Committee on Nomenclature and Statistics: Diagnostic and Statistical Manual. Washington, D.C., American Psychiatric Association, Mental Hospital Service, 1952.
- (2) American Public Health Association Program Area Committee on Mental Health: "Mental

Disorders: A Guide to Control Methods." American Public Health Association, Inc., New York, 1962.

- (3) ASHER, RICHARD. "Munchausen's Syndrome." *Lancet*, Feb. 10, 1951, 339-341.
- (4) BARKER, J. C. "The Syndrome of Hospital Addiction (Munchausen Syndrome)." *Journal of Mental Science*, 108 (1962), 167-182.
- (5) BARTON, RUSSELL. "Institutional Neurosis." Bristol, John Wright & Sons, Ltd., 1959.
- (6) BARTON, RUSSELL. "Consideration, Clinical Features and Differential Diagnosis of Institutional Neurosis." *Proceedings of the Third World Congress of Psychiatry*. 2 (1961), 890-895.
- (7) BIRCH, HERBERT G. (ed.): "Brain Damage in Children/The Biological and Social Aspects." Baltimore, The Williams & Wilkins Co., 1964.
- (8) BLEULER, E. "Das autistisch-undisziplinierte Denken in der Medizin und seine Ueberwindung." Berlin, Julius Springer, 1927. pp. 61 f.
- (9) CONNOLLY, JOHN. "The Indications of Insanity." Originally published in 1830. Now reprinted. London, Dawsons, 1964.
- (10) *Durham v. U.S.*, 214 Fed. 2d 862, 94 App. D.C. 228. See Overholser, Winfred. "Major Principles of Forensic Psychiatry." In Arieti, Silvano (editor). *American Handbook of Psychiatry*. New York. Basic Books, 1959. Ch. 95, pp. 1887-1901.
- (11) FREUD, SIGMUND. "Analysis of a Phobia in a Five-Year-Old Boy," (1909). In *Collected Papers*, vol. III. London, Hogarth Press and Institute of Psychoanalysis. pp. 149-289.
- (12) FREUD, SIGMUND. "Psychoanalytic Notes Upon an Autobiographical Account of a Case of Paranoia (Dementia Paranoides)" (1911). In *Collected Papers*, vol. III. London, Hogarth Press and Institute of Psychoanalysis. pp. 387-470.
- (13) FREUD, SIGMUND. "Certain Neurotic Mechanisms in Jealousy, Paranoia and Homosexuality" (1922). In *Collected Papers*, vol. II. London, Hogarth Press and Institute of Psychoanalysis. pp. 232-243.
- (14) GREGG, J. R. "The Language of Taxonomy." New York, Columbia University Press, 1954.
- (15) GRUENBERG, ERNEST M. "Epidemiology." In STEVENS, HARVEY A., and HEBER, RICK. "Mental Retardation: A Review of Research." Chicago and London, University of Chicago Press, 1964. 259-306.
- (16) GRUENBERG, ERNEST M., and BELLIN, SEYMOUR S. "The Impact of Mental Disease on Society." In LEIGHTON, ALEXANDER H., CLAUSEN, JOHN A., and WILSON, ROBERT N. (editors): "Explorations in Social Psychiatry." New York, Basic Books, Inc., 1957. pp. 341-364. GRUENBERG, E. M.: "Social Breakdown Syndrome—Some Origins." *Am. J. Psychiatry* (in press).
- (17) HARE, E. H. "Masturbatory Insanity: The History of an Idea." *Journal of Mental Science*, 108 (1962), 1-25.

(18) HINSIE, LELAND E., and SHATZKY, JACOB. "Psychiatric Dictionary, With Encyclopedia Treatment of Modern Terms." New York, Oxford University Press, 1940. 2d edition, with supplement. New York, Oxford University Press, 1953. 3d edition, HINSIE, LELAND E., and CAMPBELL, ROBERT JEAN. New York, Oxford University Press, 1960.

(19) JAMES, WILLIAM. "Principles of Psychology." New York, Henry Holt & Co., 1890. Reprinted. New York, Dover Publications, Inc., 1950.

(20) LEWIS, AUBREY. "Melancholia: A Clinical Survey of Depressive States." *Journal of Mental Science*, 80 (1934), 277-378.

(21) OVERHOLSER, WINFRED: "Major Principles of Forensic Psychiatry." In ARIETI, SILVANO (editor). *American Handbook of Psychiatry*. New York, Basic Books, 1959. ch. 95, pp. 1887-1901.

(22) MENNINGER, KARL, with MAYMAN, MARTIN, and PRUYSER, PAUL. "The Vital Balance." New York. Viking Press, 1963.

(23) MERRITT, H. HOUSTON, and PUTNAM, TRACY J. "Sodium Diphenyl Antoinate in the Treatment of Convulsive Disorders." *Journal of the American Medical Association*, 111 (1938), 1068-1073.

(24) MEYER, ADOLF. "Aims and Meaning of Psychiatric Diagnosis." Originally in *American Journal of Insanity*, 74 (1917-18), 163-168. Reprinted in Winters, Eunice E. "Collected Papers of Adolf Meyer." Vol. III: Medical Teaching. Baltimore, Johns Hopkins Press, 1951. pp. 48-51.

(25) MEYER, ADOLF. "Genetic-Dynamic Psychology versus Nosology." Originally in *Zeitschrift für die gesamte Neurologie und Psychiatrie*. 101 (1926), 406-427. Reprinted in Winters, Eunice E. (editor). "Collected Papers of Adolf Meyer," Vol. III: Medical Teaching. pp. 57-73.

(26) MEYER, ADOLF. "Leading Concepts of Psychobiology (Ergasiology) and of Psychiatry (Ergasiatry), 1938." Reprinted in Winters, Eunice E., op. cit., pp. 271-314. Especially pp. 293 ff.

(27) OGDEN, C. K., and RICHARDS, I. A. "The Meaning of Meaning." New York, Harcourt, Brace & Co., London, Kegan Paul, 1930.

(28) SIMON, MELLY; WIEHL, DOROTHY G.; BERRY, KATHERINE; and GRUENBERG, ERNEST M. "Inquiries to a Mental Health Association Concerning Treatment Facilities." *Milbank Memorial Fund Quarterly* 28 (1960), 301-361.

(29) TUKE, DANIEL H. "A Dictionary of Psychological Medicine." Philadelphia, P. Blakiston, Son & Co., 1892.

(30) United States of America. War Department Technical Bulletin TB MED 203 "Nomenclature and Method of Recording Diagnoses." Washington, D.C., War Department, Oct. 19, 1945.

(31) WHEWELL, WILLIAM. "The Philosophy of the Inductive Sciences, Founded Upon Their History." Vol. II, book XI, ch. IV, pp. 201-212. "Of

the Collection of Facts." London, John W. Parker, 1840.

(32) World Health Organization. "International Classification of Diseases." *Manual of the International Statistical Classification of Diseases, Injuries, and Causes of Death*. 1955 Revision. Geneva, World Health Organization, 1957.

(33) ZUBIN, JOSEPH (editor). "Field Studies in the Mental Disorders." New York, Grune & Stratton, 1961.

Supplementary Comments by Dr. Gruenberg as Presented at the Conference

I have said what I could on this topic in something over 10 pages, and have been advised by friends not to try to summarize it in under 10 minutes, but to give a talk on an entirely different topic.

In what I prepared I gave a long list of special—or not-so-special types of classifications that epidemiologists always use, whether they are available in good shape or not. I simply tried to catalog the kinds of categories that we must have to conduct certain varieties of research.

I can only pick out one or two parts of that list for attention now.

Some of the main points were discussed this morning. For example, if you want to determine the relationship between the frequency of a condition and a particular environment you not only have to classify patients' conditions but you must also classify their environments, and in some respects our ability to classify environments is in even worse shape than our ability to classify conditions. I also specified that one must be able to classify cases by the time of onset of their disorders.

Some have argued here about how many dimensions one should try to use in classifying cases and disorders. On this issue I would like to offer some rules of thumb about the number of dimensions and facts that it is thought important to record about a patient at a given moment in medical

history. My first rule is that the less you know about the condition a patient is suffering from, the more facts you want to record as possibly relevant to his diagnosis. When you don't know what the basic pattern is, and what the factors are modifying that pattern, all speculations are fair. Sibling rank certainly ranks as high as maternal age at birth, both of which may, in fact, be relevant in some disorders. Then as knowledge advances, and patterns become clearer, the number of crucial facts gets smaller, and often gets to be very small.

The second rule can be illustrated by considering the fact that there was time when "lobar pneumonia" was an adequate diagnosis. It stopped being an adequate diagnosis when the type of organism became important and one wanted to specify still further.

So the first point is: as knowledge advances the number of crucial facts becomes smaller. The second point is that as more knowledge is gained the number of crucial facts gets to be larger again as technology improves. I offer that as something to help keep a sense of balance in these arguments about how many facts one needs to know about a given case, in summarizing what is often called the diagnostic label.

Now in addition I would like to point out that some of the discussion about how many different dimensions one wishes specified about a case is ambiguous. In my opinion there are two kinds of multidimensional, or multiple diagnoses, being discussed simultaneously. One has to do with the different varieties of facts about a case with a single disorder: First, some people would like to specify what is known regarding the cause (and if the cause is unknown it should be stated as a case of unknown cause). Second, we want to describe in some summary language the clinical picture that the patient presents. Third, people want to specify, if they think they know it, the locus of the disorder in the organism. Of course in the rest of medicine there is a

good bit known about organ pathology; but in a condition like hypertension one does not have a specific locus for the disorder, and in many mental disorders we also do not have a specific locus. Fourth, people would like to specify the "level of functioning." I would prefer that term to the less inclusive term "disability." A very low level of functioning we call disability, but some patients with a serious disease sometimes have a high level of functioning, and we would like to specify that, too. They also function in different dimensions, which I won't go into because of lack of time. In psychiatry we have a fifth category, in my opinion, that does not ordinarily occur in medicine, and that is "trouble." Patients cause trouble to themselves, trouble to others, and trouble to the community. And this is a very important dimension in practical work.

I see no argument against the notion that it would be useful to set up a standard set group of classes covering each of these dimensions for clinical work. In addition, treatment outlook has been mentioned together with the word "prognosis." That is a brief outline of the concept of multiple dimensions in specifying a diagnosis.

The other aspect of multiple diagnoses can be summarized by a quotation:

One thing is certain: We have to get away from the idea of "one person one disease." Where would general pathology stand if it had to conform to the reports of the so-called "cause of death" without qualifications?—*Adolf Meyer*

The notion that we have in a classification of disorders something that puts us in a position to classify each patient only one way within that classification is completely illogical. I would like to take the time I have to underline this strongly. When we discuss different kinds of classifications we must specify whether we are interested on the one hand in classifying the patients (in a hospital, or in a clinic), or, on the other hand, in classifying the ways in which

human beings get disordered in their mental functioning.

If we are talking about the classification of mental disorders, which is usually the case in research work, then the notion that a mentally retarded person cannot have a neurosis is obviously nonsense; and the notion of confining investigation into the cause, say, of a particular neurotic symptom, to those who do not have schizophrenia and are not mentally retarded, and have no brain damage, is a bit fatuous.

From the research point of view, it seems to me, we want a classification which provides a basis for characterizing the disorders which people have; we then identify a group of people as having the same pattern of disorder, whether or not each of these people have some other disorder in addition.

Of course in conducting research, studying a population of patients each of which has several disorders, introduces well-known complexities. I am not arguing against selecting a population for study that has only one disorder.

I am arguing against the notion that the same classification will work and be constructive in trying to classify disorders of human mental functioning and in classifying psychiatric patients. The two problems are different, and the labels, although often the same, do not have the same meaning or implications.

I should point out more explicitly, I suppose, since it was brought up earlier, what these remarks should have made quite clear: In spite of Dr. Henry Brill's implications, I have no intention of defending the American Psychiatric Association's "Manual" on classification. It is sometimes called a classification of mental disorders. Its instructions tell you how to use it as a classification of psychiatric patients. It has several dimensions within it that make it possible very readily to decide not to make a diagnosis of schizophrenia and mental retardation in the same patient. In fact it

makes it hard to make those two diagnoses in the same patient.

Classification of Mental Disorders For Epidemiologic and Medical Care Purposes: Current Status, Problems, and Needs

*Morton Kramer, Sc. D.*¹

1.0 Introduction

This paper will consider the nomenclature and classification of mental disorders from the point of view of the biostatistician whose task it is to provide data on the incidence and prevalence of mental disorders, patterns of utilization and effectiveness of psychiatric services, and the fate of patients with such disorders in relation to time, space, and demographic, socioeconomic, and related characteristics of patients and populations. Such data are essential to planning appropriate programs for the prevention and control of mental disorders and for carrying out research to establish the efficacy of such programs. They are also needed by epidemiologists who are attempting to uncover clues to the causes of these disorders, to explain their differential distribution in various population groups within and between countries and changes in such distributions over time.

The systematic basic data that have been used in studies of the distribution of mental disorders in populations as of specific moments in time and over time have, for the most part, originated in mental hospitals (1). More recently data have become available from outpatient psychiatric clinics, general hospitals with psychiatric services and

¹ Chief, Biometry Branch, Office of Program Planning and Evaluation, National Institute of Mental Health, Public Health Service, Department of Health, Education, and Welfare, Chevy Chase, Md.

related facilities that provide care, treatment, and rehabilitation services for persons with mental disorders. The diagnoses placed by the attending psychiatrists on patients admitted to these facilities thus constitute the basic data from which distributions of mental disorders and temporal trends in the rates of occurrence of such disorders have been established in the United States and elsewhere. Such data are medical care data and as such provide information only on that portion of the population with a specific disorder admitted to the facilities in question. Data so derived must be differentiated from morbidity data which provide measures of the total incidence and prevalence of a disease counting both those who enter into treatment and those who do not. Quite a few community surveys have been carried out to establish the prevalence of mental morbidity in various population groups in the United States and elsewhere (2). However, none of these surveys has yielded a method which is generally applicable for assessing periodically the prevalence or incidence of mental disorders as a group or of specific mental disorders. Consequently, diagnoses derived from the medical records of patients admitted to psychiatric facilities still remain the principal source of systematic data on the occurrence of specific mental disorders.

2.0 Systems of Diagnostic Nomenclature and Classification Currently in Use in the United States

The diagnostic terminology used currently by psychiatrists in the United States is that which is published in the fourth edition of the "Standard Nomenclature of Diseases and Operations" (3).² The definitions of the terms in the diagnostic nomenclature

have been spelled out in the "Diagnostic and Statistical Manual: Mental Disorders" prepared by the Committee on Nomenclature and Statistics of the American Psychiatric Association and published in 1952 (4).

In the United States the system for classifying these diagnoses for statistical purposes is based on the seventh revision of the "International Statistical Classification of Diseases, Injuries, and Causes of Death" (ICD), modified to meet some of the unique problems created by the diagnostic terminology used in this country (5). Over the past several years considerable activity has been going on within countries to assist the World Health Organization in its preparation of the eighth revision of the ICD. A proposal was adopted in July 1965 by the International Revision Conference of the World Health Organization and, following adoption by the World Health Assembly, will be submitted to governments for comment. This includes a revision of the classification of mental disorders.

3.0 Differences Between a Nomenclature and Statistical Classification of Diseases

Since the phrases "nomenclature of diseases" and "statistical classification of diseases" will be used frequently during the course of this conference, it is essential at the outset to differentiate between them. Since the "International Classification of Diseases" will also be mentioned, a brief description of it will also be given.

A nomenclature of diseases is "a list or a catalog of approved terms for describing and recording clinical and pathological observations. To serve its full function it should be extensive so that any pathological condition can be accurately recorded. As medical science advances a nomenclature must expand to include new terms necessary to record new observations (3) * * * ." The "Standard Nomenclature" is such a catalog of diseases. Designed primarily for use by clinicians, it attempts to include every

² The fifth edition of the Standard introduced certain revisions into the psychiatric nomenclature, but these revisions have not been incorporated into the APA statistical manual, and as a result, are not used in the statistical reporting in mental hospitals and outpatient psychiatric clinics. This may pose a problem as more general hospitals introduce psychiatric services and use the fifth edition of the Standard for medical-record indexing, statistics, and related matters.

disease which is clinically recognizable in a way that avoids repetition and overlapping. It does this by assigning every disease to a topographic system of the body and then to a specific etiologic classification. Mental disorders are assigned to the psychobiologic unit of the system devoted to the body as a whole and to an etiologic system appropriate to the disorder being cataloged.

The extensive and complete specificity of a nomenclature prevents it from serving satisfactorily as a statistical classification; that is, a classification designed primarily to furnish statistical data about groups of diagnostic entities. In such a classification a specific disease entity has a separate title only when its separation is warranted because of the frequency of its occurrence or because of its importance as a morbid condition. Although each disease does not constitute a separate title in a statistical classification, every disease or morbid condition must have a definite and appropriate place for inclusion in one of the categories of the statistical classification. An essential phase of the development of a statistical classification, therefore, is agreement on the number and titles of the categories of the classification and the diagnoses to be included under each such title. The statistical classification groups several related diagnostic entities into a single category and, generally, has a residual category for "all other disorders." This latter category includes all diagnostic terms in the nomenclature that, for one reason or another, cannot be assigned to a defined category of the statistical classification (6).

Thus, a statistical classification of mental disorders must be confined to a limited number of categories which will encompass the entire range of these disorders. Such a classification must be constructed in a way that makes it possible to group all of the clinical terms used in psychiatry into a limited number of well-defined, meaningful, and reasonably homogeneous categories.

4.0 The International Classification of Diseases

The "International Classification of Diseases" (ICD) has been developed essentially to provide a list of diseases for compiling statistics of causes of death and causes of morbidity throughout the world. Its widest use has been in the compilation of mortality statistics. To facilitate this, rules have been developed to select the underlying causes of death in instances where multiple illnesses are listed on the death certificate. No similar rules have been developed for presentation of morbidity statistics. The classification is being adapted increasingly for the purposes of indexing hospital records (7).

Since the ICD is designed for general use in the study of disease, it is eclectic in its scheme of classification. It deals first with the diseases caused by well-defined infective agents; these are followed by categories for neoplasms, allergic, endocrine, metabolic, and nutritional diseases. Most of the remaining diseases are arranged according to the principal anatomical site with special sections for mental disorders, complications of pregnancy and childbirth, certain diseases of early infancy, senility, and ill-defined conditions, including symptoms. Its final section provides a dual classification of injuries according to external causes giving rise to injury and the nature of injury.

The classification of mental disorders of the ICD (sec. V) that has been available since 1948 has not been widely used by psychiatrists in different countries (8). It is quite inadequate for classifying the diagnostic terms of the 1952 revised psychiatric nomenclature of the American Psychiatric Association. The reason for this was that the ICD did not provide categories for coding certain new diagnostic categories that were introduced in this revision (particularly chronic brain syndromes with neurotic or behavioral reactions). In 1960 a subcommittee on classification of mental disorders of the U.S. National Committee

on Vital and Health Statistics was appointed to develop a classification to correct this situation. This subcommittee worked closely with a corresponding committee of the United Kingdom to develop a single classification for classifying for statistical purposes the diagnostic terms used in clinical psychiatry in the United States and United Kingdom. A series of recommendations were also made to the WHO by national committees from Norway, Denmark, France, U.S.S.R., and other countries for consideration by the Expert Committee on Health Statistics of the WHO. As a result of deliberations at several meetings of this Expert Committee between 1961 and 1964 compromise solutions were worked out to provide a statistical classification which could be used to code and classify the diagnostic terms used in clinical psychiatry in different countries of the world for the production of reasonably comparable statistical tabulations on psychiatric disorders.

Over the years many criticisms have been leveled against statistical classification of mental disorders, both that in the ICD and its modified form in use in the United States. Many of these criticisms stem from failure of persons who use these classifications and the statistics derived therefrom to understand the underlying purpose of the classification. The classification currently in use in the United States is intended to classify the diagnostic terms used in clinical psychiatry for purposes of providing statistics on the diagnostic classification of patients using various types of psychiatric facilities, for classification of psychiatric cases detected in morbidity surveys, for studies of fate of specific diagnostic groups, and for indexing records of patients under care in psychiatric facilities. Criticisms are frequently leveled against diagnostic distribution of patients admitted to hospitals or outpatient facilities on the grounds that these distributions tell us little about the patient population, they do not describe patients' presenting problems, they do not accurately reflect the premorbid personality

of patients admitted, they do not reflect the condition or circumstances that precipitated the current attack of mental disorders. The fact that such questions are not answered by the statistical classification of mental disorders is not the fault of the classification since it was not created to answer such questions. A fundamental principle of classification is that the type of classification of disease to be used depends on the purpose of the study, the questions to be answered and the nature of the data to be classified. Thus, if one wants to answer questions concerning presenting problems of patients admitted to a facility, then a classification must be developed for this purpose. The same is true in relation to answering questions on symptom patterns, premorbid personality structure of patients, precipitating events, et cetera.

5.0 Reliability of Diagnosis—A Key Issue

The critical problem confronting the biostatistician who uses diagnostic data from psychiatric records is that of observer variability; i.e., the reliability and accuracy of the diagnostic terms recorded and the consistency with which an individual clinician or groups of clinicians use diagnostic terms and apply a given diagnostic term to patients with the same clinical characteristics or profiles. As Reid has pointed out (9), there are three kinds of observer variability: "consistent bias which is a reflection of the type and intensity of the training which the observer may have had; bias which is personal to himself; and the inconsistencies in his own judgment over a period of time which may be either erratic or regular."

So that diagnostic data derived from clinical records will be reliable and meaningful, it is essential not only that diagnostic nomenclature be standardized but also that standards for each diagnostic entity be developed and used by clinicians in a uniform way from time to time and place to place. If the problem is to determine whether there is a difference in the prevalence of schizo-

phrenia in white males, aged 25–44 years in community *A* as compared to community *B*, it is essential that the same diagnostic procedures and criteria be used by the psychiatrists in each area and applied by them in a standardized fashion. Unless this is done it is impossible to determine whether the differences (or for that matter the absence of differences) reflect differences among psychiatrists in their use of diagnostic criteria and/or diagnostic terms or true differences in the rates of schizophrenia. If the problem is to determine whether there is a change in the incidence rate for a specific disorder over time, then there is a need to demonstrate that whatever change has taken place between two points in time is the result of change in the rate at which the disorder has been occurring and not of differences in the criteria utilized to diagnose the disorder or in the way the observers who are collecting the data are applying the criteria.

Similarly, comparisons of the diagnostic distribution of patients admitted to mental hospitals with that of patients admitted to outpatient psychiatric clinics can be meaningful only if the same criteria are used to arrive at specific diagnoses for patients admitted to each set of the facilities. In explaining differences that occur in rates of utilization of separate facilities specific for diagnosis, the statistician must be aware not only of the selective factors that determine who is admitted to a specific facility, but also of the differences in training and orientation of the psychiatrists who work in the different settings and the possible effect this may have on their diagnostic habits.

Similar considerations go into comparisons of the evaluation of the effectiveness of therapies or therapeutic programs on specific diagnostic groups of patients.

The biostatistician is also constantly faced with the problem of deciding how fine or, conversely, how condensed diagnostic categories within a statistical classification should be made in order to present mean-

ingful tabulations of morbidity data. This is particularly important in the psychiatric field where rigid diagnostic criteria do not exist, diagnostic terms and concepts have been used inconsistently and have changed over time, and frequently it may be difficult to differentiate precisely between different diagnoses within a major category. Thus, it may be more reliable to present data on the functional psychoses as a group rather than on the separate diagnostic categories that comprise it. The same considerations apply to presenting data for disorders associated with alcoholism as a group as opposed to presenting separately data for acute and chronic brain syndromes associated with alcoholism and alcoholic addiction; or for disorders of the senium rather than separate data for psychosis with cerebral arteriosclerosis and senile psychosis.

The grouping used depends on the reliability and consistency with which diagnostic terms have been used. Thus, if the statistician has been working in a hospital where there has been little turnover in clinical staff over the years, and there is evidence that the diagnostic methods have remained reasonably consistent over time, his selection of diagnostic categories for presentation of statistics presents no major problem. Problems may arise from the size of the number of cases that fall within a given category and the limited number of cross tabulations by age, sex, marital status, residence, etc., that may be made from the data. On the other hand, the statistician working at the city or county, State, national, or at the international level faces quite different problems in making decisions on the grouping of the data. Here, the problem is not one of the number of cases in a given diagnostic category, but of the consistency and reliability with which diagnostic standards have been applied and with which certain diagnostic terms are used to characterize individuals who present the same clinical picture. The problem is considerably exacerbated in international comparisons where the psychiatrists in one

country use a diagnostic term derived from a theoretical concept that is not acceptable, or at least not used, by psychiatrists in another country. Cases in point are the use in the United States of the acute and chronic brain syndrome, a concept with a set of diagnostic terms which are not used in many other countries of the world, and the diagnosis of *bouffée delirante* used in France and that of *psychogenic psychosis* used in Norway for which there are no precise counterparts in the U.S. scheme of diagnostic terminology.

6.0 Examples of Differences in Diagnostic Distributions

The problem of interpretation of differences in diagnostic distributions may be illustrated by several tabulations of diagnostic data. The first, shows variations in time trends in first admission rates to State

mental hospitals in the United States by diagnosis; the second, variations between different types of psychiatric facilities in diagnostic distributions of admission; and the third, a comparison of first admission rates to mental hospitals in the United States and England and Wales.

6.1 Trends in first admission rates to State and county mental hospitals.—Table I shows the age-adjusted first admission rates specific for diagnoses for the years 1940, 1950, and 1960.

Striking changes have been taking place. Perhaps the only diagnostic category for which diagnostic reliability has been reasonably satisfactory over the years is “brain syndromes associated with syphilis.” For a variety of reasons quite well known to this conference, the decrease in this admission rate can be accepted as real. The large increases in the rates for the psychoneuroses, personality disorders, and alco-

TABLE I.—Age-Adjusted¹ 1st Admission Rates and Percent Change to Public Mental Hospitals 1940, 1950, and 1960 for Selected Diagnoses

Diagnosis	Rates per 100,000			Percent change		
	1940	1950	1960	1940-50	1950-60	1940-60
All mental disorders	90.2	94.6	101.3	4.9	7.1	12.3
Brain syndromes	42.3	38.8	25.4	-8.3	-34.5	-40.0
Diseases of the senium	22.0	24.5	19.0	11.4	-22.5	-13.6
Syphilitic	7.6	3.2	.6	-57.9	-81.3	-92.1
Other (excluding alcoholics)	12.7	11.1	5.8	-12.6	-47.8	-54.3
Functional psychoses	30.0	32.6	33.7	8.7	3.4	12.3
Schizophrenics	17.2	22.0	25.3	27.9	15.0	47.1
Affective (including involuntions)	11.2	9.5	7.4	-15.2	-22.1	-33.9
Other	1.6	1.0	1.0	-37.5	0	-37.5
Disorders associated with alcoholism	9.3	11.4	15.0	22.6	31.6	61.3
Brain syndromes	4.8	4.2	5.4	-12.5	28.6	12.5
Addiction	4.5	7.2	9.6	60.0	33.3	113.3
Psychoneurosis	2.9	4.7	7.7	62.1	63.8	165.5
Personality disorders	1.8	2.8	10.5	55.6	275.0	483.3
Mental deficiency	1.4	1.2	3.1	-14.3	158.3	121.4
All other	2.5	3.1	5.8	24.0	87.1	132.0

¹ U.S. 1950 civilian population 10 years of age and over was used as a standard.
Source: U.S. Department of Health, Education, and Welfare, PHS, NIMH; "Patients in Mental Institutions, Part II, 1940, 1950, 1960"; "State and County Mental Hospitals."
Prepared by: Hospital Studies Section, Biometrics Branch, National Institute of Mental Health, October 1964.

holic addictions raise questions as to whether they represent true increases in rates of admission for these diagnostic groups or differences in diagnostic fads and criteria or other factors that lead mental hospital psychiatrists to place nonpsychotic diagnoses on increasing numbers of patients.

The increase in the first admission rate for schizophrenia and decreases in that for affective psychoses are also difficult to evaluate, not only because of community factors that can affect the level of these rates, but also because of inconsistencies and differences in the way clinicians have used these diagnoses over the years. As a result, the statistician also presents a trend for the group of functional psychotic disorders, since he believes the trend for the group of disorders rather than for the individual disorders within the group may be more likely to provide a more accurate account of what is going on.

However, administrators, clinicians, and others persist in asking questions about the trend in the admission rate for the population of patients diagnosed as schizophrenic, and of that diagnosed as manic depressive. The reasons for such questions are many, stemming from the need to justify a budget, to plan programs, to evaluate efforts to treat more effectively specific categories of patients and to determine the effect of community facilities on changing the patterns of admissions to mental hospitals for specific diagnostic groups. For example, considerable effort has been expended to change the patterns of care of the schizophrenics, and questions are raised frequently about the effect of these efforts on the admission rates for patients with this diagnosis. Much interest is being expressed in developing programs for the prevention and control of depression. This leads to questions concerning the reasons for the decrease in admissions of affective psychoses to mental hospitals.

6.2 *Differences in diagnostic distribution of patients admitted to various com-*

munity facilities.—To throw more light on the use of various community facilities for various types of patients, psychiatric case registers have been established in several areas (10). This mechanism facilitates the collection and linkage of reports on individual patients admitted to all of the psychiatric facilities that serve the residents of a defined geographic area. It is then possible to study patterns of utilization of these services on a person basis rather than on an event basis, specific for diagnosis, age, sex, and various demographic factors.

Table II presents a tabulation of admission rates by diagnosis to the entire universe of facilities serving residents of Maryland, and to three broad groups of these facilities: (a) State and county mental hospitals; (b) private mental hospitals, psychiatric services in general hospitals, and VA hospitals; and (c) outpatient clinics.

This tabulation was prepared to investigate the utilization of facilities by persons with depressive disorders. The register data could be tabulated in greater diagnostic detail than that reported to the NIMH nationally so that it is possible to present admission rates for depressive disorders in the following detail:

Psychotic Disorders — Affective disorders: Involutional psychotic reactions; manic-depressive reactions; psychotic-depressive reactions.

Psychoneurotic Disorders—Depressive reactions.

It is quite clear from table II that the admission rates for affective psychoses and psychoneurotic depressive reaction are markedly different among the three groups of facilities. Thus, the admission rate for the affective psychoses is highest for the State and county mental hospitals and that for the psychoneurotic depressive reactions is considerably higher for the other types of facilities.

A major concern here relates to whether diagnostic criteria and terminology are used in a comparable fashion by psychi-

TABLE II.—Age Adjusted Admission Rates¹ Per 100,000 Population for Selected Diagnoses, by Type of Facility

Diagnosis ¹	Admissions July 1, 1962–June 30, 1963			
	Total ³	Type of facility		
		State and county mental hospitals ⁴	Private, general, VA hospitals	Outpatient psychiatric clinics
TOTAL MARYLAND				
All patients	546.0	228.4	101.4	296.5
Schizophrenic reactions (22)	116.8	69.7	21.8	57.4
Affective psychoses, total	18.2	8.7	6.9	4.6
Involutional psychotic reaction (20.0)	1.5	1.5	(—)	(*)
Manic-depressive reactions (21.0, 21.1, 21.2) . .	8.2	4.0	2.5	2.6
Psychotic-depressive reaction (21.3)	8.4	3.1	4.4	1.9
Psychoneurotic-depressive reaction (40.5)	59.2	12.6	27.4	25.2
Total depressive reactions (20, 21, 40.5)	77.5	21.3	34.6	29.9
Diseases of the senium (15.0, 17.1, 17.2)	33.6	29.0	3.9	4.1
Alcoholic disorders (02.1, 13.0, 52.3)	67.2	53.2	9.6	15.3
All other	250.9	55.2	31.5	189.8
All other diagnoses	201.2	48.9	19.6	155.6
Undiagnosed	49.7	6.3	11.9	34.2
BALTIMORE CITY				
All patients	771.0	390.2	75.8	442.2
Schizophrenic reactions (22)	204.8	131.8	20.3	111.8
Affective psychoses, total	25.3	13.6	7.1	8.2
Involutional psychotic reaction (20.0)	2.3	2.1	(—)	(*)
Manic-depressive reactions (21.0, 21.1, 21.2) . .	11.2	6.3	1.8	4.7
Psychotic-depressive reaction (21.3)	11.8	5.2	5.3	3.4
Psychoneurotic-depressive reaction (40.5)	70.8	19.4	19.2	41.6
Total depressive reactions (20, 21, 40.5)	96.1	33.0	26.2	49.8
Diseases of the senium (15.0, 17.1, 17.2)	35.8	30.3	3.6	5.8
Alcoholic disorders (02.1, 13.0, 52.3)	122.2	108.8	3.3	31.4
All other	312.1	86.3	22.4	243.4
All other diagnoses	241.7	76.2	16.1	185.2
Undiagnosed	70.4	10.1	6.3	58.2

¹ Adjusted to 1960 U.S. civilian population total age distribution.

² The numbers in parentheses following the separate diagnostic categories are code numbers utilized in the statistical classification of the 1952 "Diagnostic and Statistical Manual, Mental Disorders" of the APA. See pages 78–86 of reference 4.

³ This total is less than the sum obtained from adding the rates for each group of facilities. The reason for this is that patients were admitted to more than 1 type of facility. The diagnosis used for admissions to the total universe of facilities is that which was placed on the patient at the 1st facility to which he was admitted. The diagnosis utilized in the distribution for the separate types of facilities is that which was placed on the patient during his 1st admission to that type of facility.

⁴ Including returns from long-term leave.

NOTE.—Rates not shown for less than 5 patients (—)for 0 entry; (*) for 1–4 entry).

atrists working in State mental hospitals and those working in private hospitals, general hospitals with psychiatric services, and outpatient clinics. The fact that the rate of admission to the State mental hospitals in Maryland for psychoneurotic depressive reactions exceeds that for affective psychotic disorders as a group raises further

questions about the consistency with which diagnostic criteria are utilized within the State hospitals with respect to diagnosis of depressive disorders. Indeed, in a recent study of diagnostic consistency of patients who received services in more than one psychiatric facility in Monroe County (N.Y.) it was demonstrated that there is a

marked disagreement among psychiatrists in distinguishing psychotic from non-psychotic-depressive disorders (11).

At this point, the best that can be done is to present the data as reported. It must be emphasized, however, that the meaning of data on the individual titles in the depressive disorders and of the inferences that can be drawn from them concerning differential patterns of admission and of subsequent experience of specific groups of patients are only as good as the reliability and validity of the original diagnostic data.

6.3 International comparisons of diagnostic distributions.—The increasing interest in cross-cultural comparisons of rates of mental disorders and in the types of programs that have been developed in various countries of the world to treat and rehabilitate the mentally ill has increased the demand for comparative data on the mental disorders at the international level. Here, the problems of developing reliable comparative diagnostic data are even more complicated than those attendant to developing such comparisons within a country. As in the United States, the systematically collected data on the mentally ill in most countries of the world (as opposed to one-time prevalence surveys) are derived essentially from the records of mental hospitals.

Kramer has called attention to the extraordinary differences between the patterns of first admissions specific for diagnosis to the mental hospitals of England and Wales as compared to the United States (12). The shapes of the first admission curves for all diagnoses are the same (fig. 1). However, there are striking differences in the rates for different diagnostic groups. Very large differences exist between the first admission rates for manic-depressives and cerebral arteriosclerotic psychoses (figs. 2 and 3). Thus, the rate for manic-depressives in England and Wales is about 18 times the rate to public mental hospitals in the United States and 9 times the rate to our

public and private mental hospitals combined. In addition, the age distribution for the manic-depressive rates in England and Wales is quite different than that in the United States. The rate in England and Wales is at a very high level in the age groups 55 and over and is at a very low level in the United States. In the age group 55–64 years where the maximum rate occurs in England and Wales, its rate is about 11 times the rate to the U.S. public and private hospitals combined. Another point of interest is that the rates for manic-depressive psychosis in England and Wales have continued to increase over the years, whereas those in the United States have continued to decrease (13). Even when all functional psychoses are combined (fig. 4) the admission rates to the England and Wales hospitals are still extraordinarily high, particularly in the older age groups.

At this time, it cannot be determined how much of these differences are attributable to (a) differences in the diagnostic criteria and diagnostic nomenclature used by English and American psychiatrists to characterize patients who may present the same clinical picture at time of admission, and (b) differences in the clinical characteristics of patients admitted to the mental hospitals in England and Wales and in the various States of the United States. Systematic studies are needed of differences in the underlying criteria used by British and United States psychiatrists in arriving at a particular diagnosis and of the differences that would occur in the diagnostic distributions of patients admitted to various types of psychiatric facilities in each country if each of the patients was diagnosed independently by a representative group of American psychiatrists and by a representative group of English psychiatrists. A group of investigators from the New York State Psychiatric Institute and from the Maudsley Hospital and Medical Research Council (MRC) Social Psychiatry Research Unit in London are

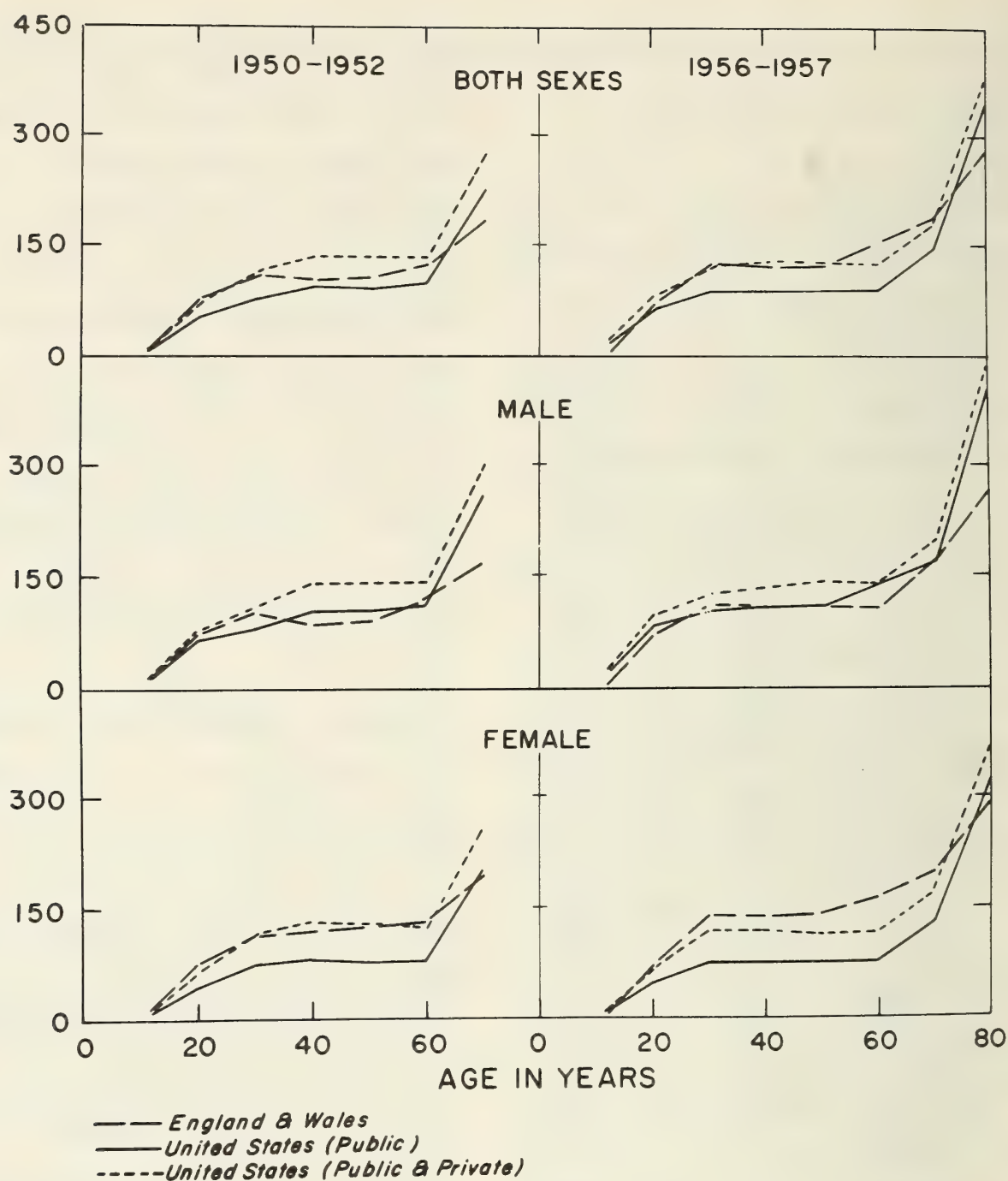


FIGURE 1.—Age-sex specific first admission rates per 100,000 population to public mental hospitals in the United States, 1950, 1957, and in England and Wales, 1952, 1956, and to public and private mental hospitals combined in the United States.

planning investigations to resolve some of these problems.

Another investigation is suggested by this comparison. Of particular importance in the effective utilization of drugs (or other therapies) that have specific effects is the need to diagnose accurately patients that have the specific disorders and symptoms that respond to such therapies. If patients are being inaccurately diagnosed and classified the possibility exists that they may be treated improperly or at least inadequately.

The large differences between England and Wales and the United States in the

rates of first admission from the population 65 years and over for affective and organic psychoses, taken in conjunction with findings reported by Roth concerning differences in outcome of hospitalization for these two broad categories of patients, raise important questions. Roth emphasizes the importance of not considering the mental disorders of old age as a single entity and the need to make a careful differential diagnosis (14). He states that where affective disorders "are regarded as early organic psychoses, valuable opportunities for effective treatment may be missed." It would seem important to deter-

mine to what extent, if at all, affective psychoses in old age in the United States are being regarded as organic psychoses and valuable treatment opportunities are being missed because of this.

7.0 Needs

7.1 *Studies of observer reliability.*—As indicated earlier, the major need for assuring reliable diagnostic data on the mental disorders is the development of diagnostic methods and instruments that would lead to greater precision and consistency in the way psychiatrists utilize diagnostic terms. Such methods are needed

in order to minimize the effect of observer variation in producing differences in diagnostic distributions between groups of psychiatric patients from place to place and time to time.

Studies have been undertaken of observer variation on the basis of data derived from independent diagnostic assessments of patient records. Although such a method provides a simple way of providing independent observers with a written clinical record for assessment, it has several shortcomings because of the many factors that may affect the completeness and accuracy of data contained in such records. Foremost among these factors are variability in

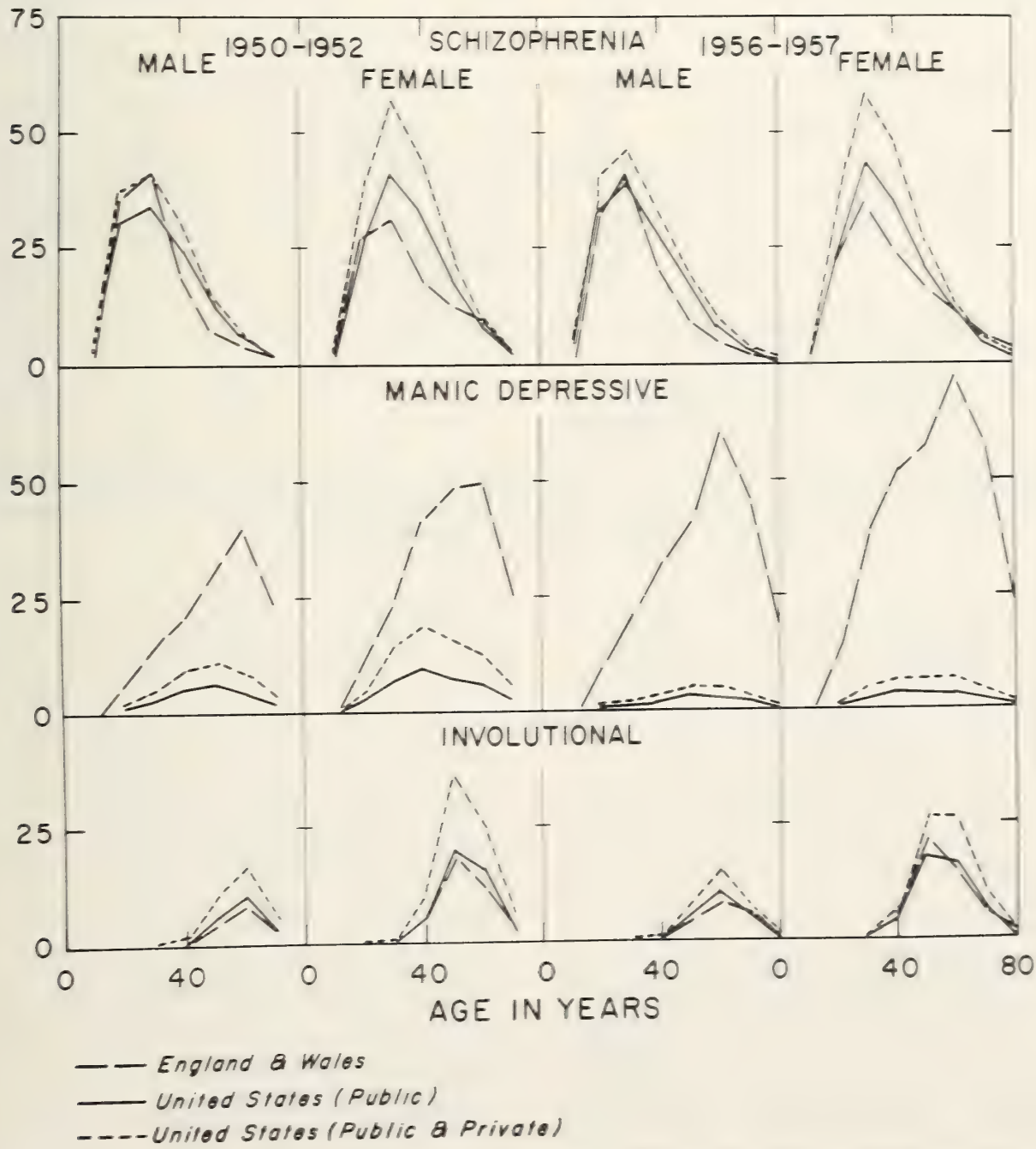


FIGURE 2.—Age-sex specific first admission rates per 100,000 population to public mental hospitals in the United States, 1950, 1957, and in England and Wales, 1952, 1956, and to public and private mental hospitals combined in the United States.

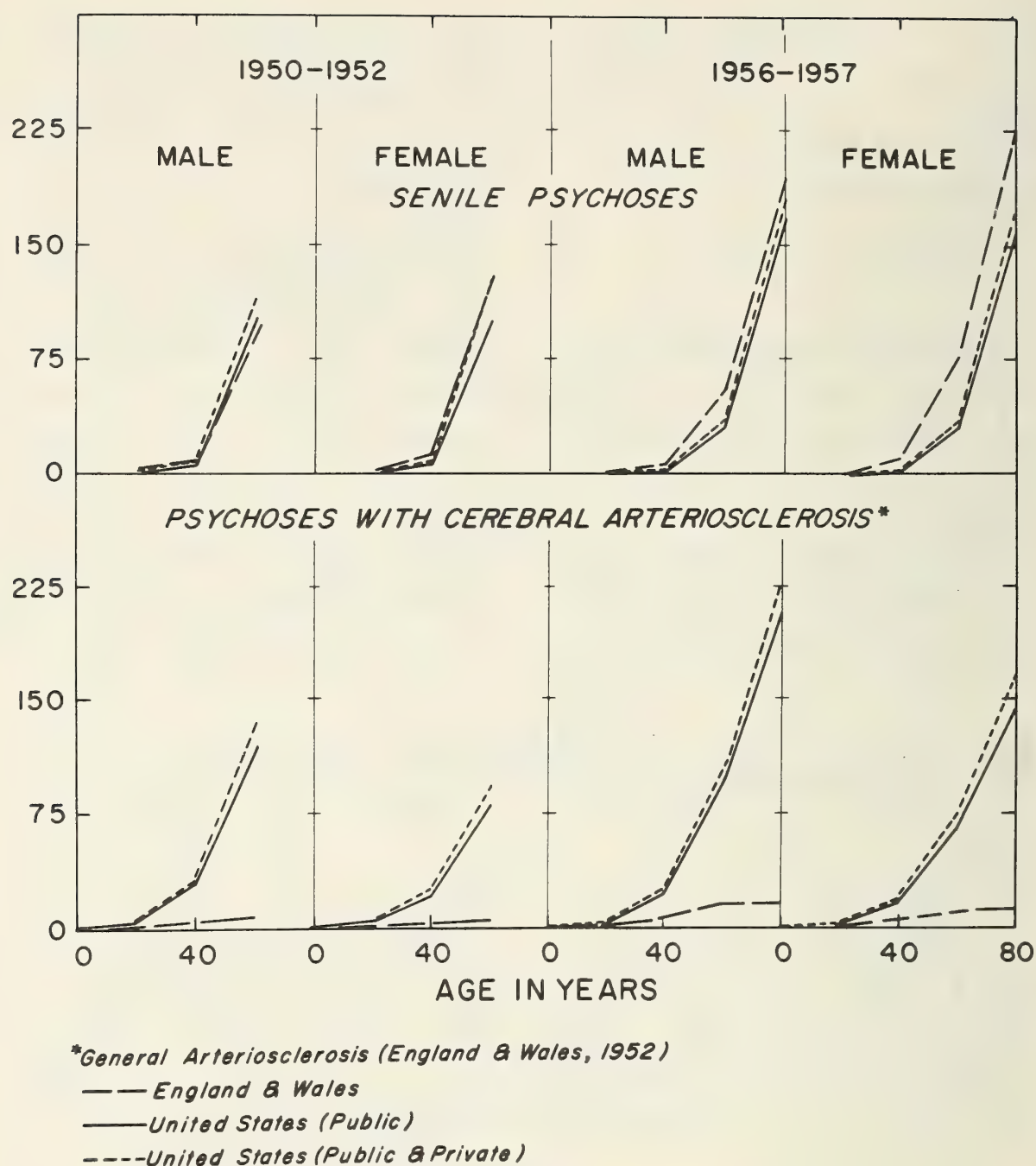


FIGURE 3.—Age-sex specific first admission rates per 100,000 population to public mental hospitals in the United States, 1950, 1957, and in England and Wales, 1952, 1956, and to public and private mental hospitals combined in the United States.

the skill of the original examiner, his thoroughness in eliciting pertinent data and the accuracy and completeness with which he entered the observations in the clinical record. Further, the psychiatrist who is reviewing the study records has not observed an actual face to face interview between patient and clinician.

What is needed to get at the problem of observer variability is a method for allowing a set of observers to arrive at an independent assessment of the same patient. These ratings must be based on identical anamnestic and clinical data and a standardized clinical interview. The essential tools for such research would include:

Standardized anamnesis; social adaptation history; clinical examination and clinical interview; a video tape recording of a live clinical interview to produce an audio-visual record which could be shown to as many independent observers as desired, a standardized method for recording the observations and scoring them. Profiles could be developed for individual patients and groups of patients with the same diagnostic labels and analyzed by various characteristics of diagnosticians as well as patients. The variability of profiles within and between given diagnostic groups could be determined as well as variability of diagnostic groups within and between cate-

gories established on the basis of similarity of profiles. Sufficient work has been done with the video taping of diagnostic interviews to demonstrate that this is an entirely feasible method of recording a clinical interview. Considerable progress has been made in the statistical aspects of profile analysis so that methods are available for this aspect of the research. For example, methods have been developed which make it possible to determine similarity and differences in the levels of profiles of patients (15). Also, considerable progress has been made in developing a mental status schedule for simultaneously examining, recording, and evaluating the mental status of a psychiatric patient by means of a standard interview (16).

The research outlined above should yield instruments which would tend to minimize variability resulting from differences in the way data on patient characteristics are collected and recorded. Their use would make it possible to determine the extent to which there is agreement and disagreement among psychiatrists within as well as between different countries on diagnoses placed on identical sets of patients, when identical data are used as the basis for the diagnostic assessment. It would also be possible to determine the profiles of the patients on which there are various degrees of observer agreement. The results of such research offer the possibility of classifying patients in a more objective and uniform

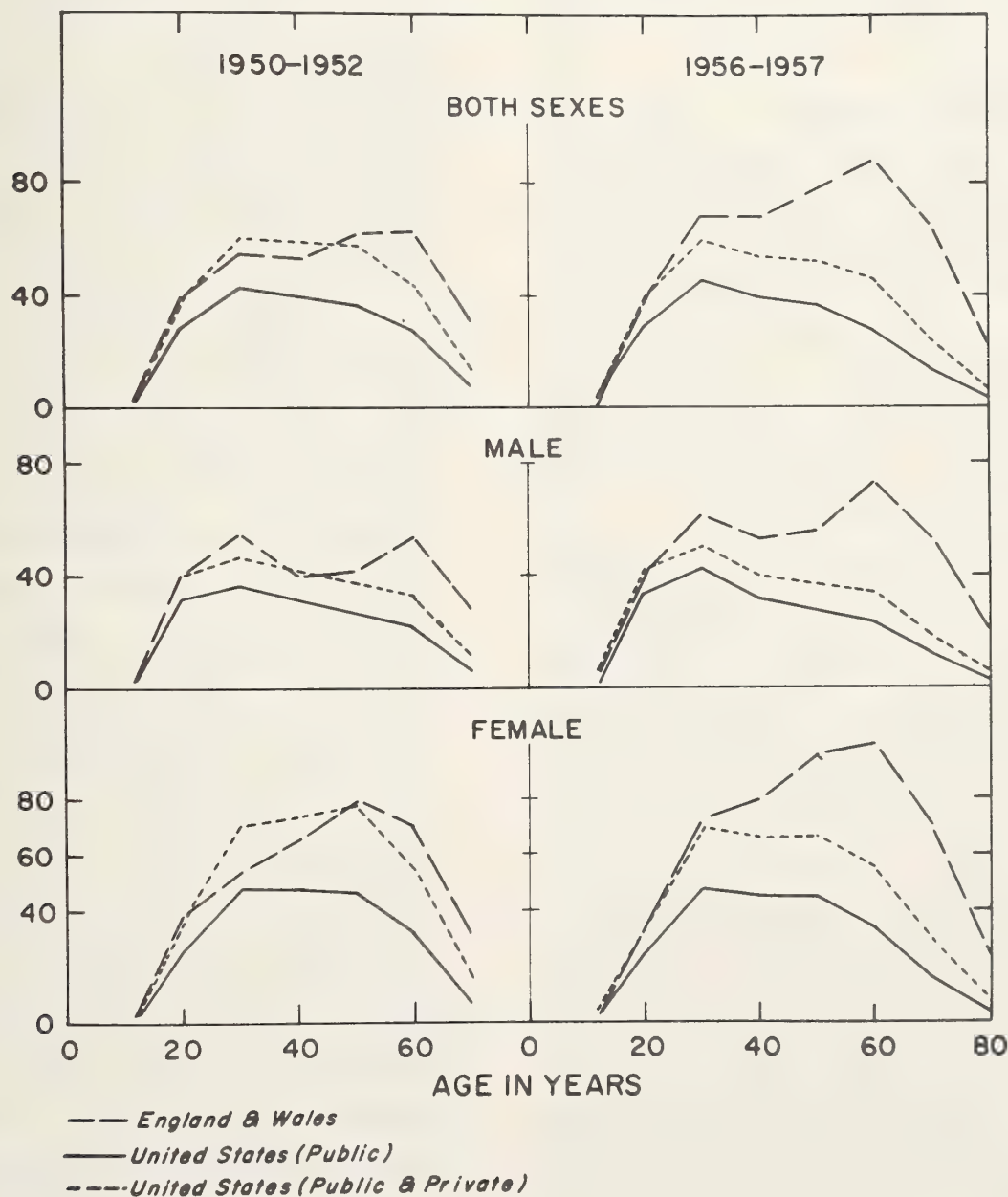


FIGURE 4.—Age-sex specific first admission rates per 100,000 population to public mental hospitals in the United States, 1950, 1957, and in England and Wales, 1952, 1956, and to public and private mental hospitals combined in the United States, 1950, 1957: functional psychoses.

fashion—either on the basis of diagnosis or similarity of profiles.

It should be emphasized that such research is quite complex. It will take time to bring it to a point where practical applications could be made of the techniques so developed for classifying patients. Indeed, it is quite possible that the methods derived from such research would be applicable only in certain clinical or research settings and projects (for example, the evaluation of a specific type of drug therapy, or a well-defined study of prevalence). In the meantime methods are needed for classifying patients by diagnosis for planning services, for treatment, and many other purposes. Certain practical steps can be taken to obtain greater uniformity in psychiatric diagnostic practice and usage of diagnostic terminology. Some suggestions follow.

7.2 Diagnostic manuals. — The APA Manual has been widely used in hospitals and clinics throughout the United States. This will probably be revised to bring it up-to-date and to modify it in ways that will relate it more clearly to the revised “International Classification of Mental Disorders.” Regional and local workshops should be held to instruct clinicians in the use of this revised manual.

The United Kingdom Subcommittee on Classification of Psychiatric Disorders is also preparing a glossary of psychiatric disorders. It would be most encouraging for international work if a working party from both our countries could reach agreement on a single diagnostic manual and glossary of diagnostic terms that could be utilized in both the United States and United Kingdom. The WHO has urged all countries to develop glossaries giving operational definitions for the categories and subcategories of the section on “Mental Disorders” of the revised ICD. The WHO is also developing a research program to provide a basis for obtaining greater uni-

formity in the usage of psychiatric diagnostic terminology throughout the world.

7.3 Medical records committees at the institutional and community level.—The medical record is the basic source of information for planning patient care, in providing a means of communication between the physician and other professional groups contributing to the patient's care and furnishing documentary evidence of the course of the patient's illness and treatment. It also serves as a basis for review, study and evaluation of the medical care rendered the patient. These records also serve as the source document for compilation of medical statistics. All hospitals accredited by the Joint Committee on Accreditation have medical records committees which periodically review the medical records from a clinical standpoint, to evaluate the quality of care given the patient and the completeness as well as the accuracy of these records (17). These committees have an opportunity to study the consistency with which staff utilizes diagnostic terms to characterize patients with similar clinical profiles. With the increasing numbers of mental hospitals that will apply for accreditation by the Joint Committee to qualify for benefits under the new Medicare legislation, more hospitals will be establishing and utilizing such committees within their own institutions.

The developing community mental health centers program provides an opportunity for extending the concept and functions of the medical records committees to all of the psychiatric facilities in an area. Under the regulations governing the development of plans for comprehensive centers (18), any one of the collaborating elements of inpatient, outpatient, intermediary, and related services that comprise the comprehensive program and accept transfers of patients from any other element can exchange clinical information on patients so transferred. A records review committee, established to accomplish on a communitywide basis what such

a committee does on an individual institution basis, should lead to more effective utilization and provision of services provided by each participating facility. Thus, a committee consisting of representatives from psychiatric services in general hospitals, public and private mental hospitals, outpatient clinics, day care centers, etc., could review periodically samples of records from each facility to determine the completeness of the data recorded, the diagnostic formulations, and the appropriateness of treatment programs and of various transfer actions, etc. It would also be in a strategic position to develop procedures for obtaining greater uniformity in diagnostic and clinical practice and in use of diagnostic terminology among the psychiatrists in the participating facilities.

7.4 *Classifications of aspects of psychiatric morbidity other than diagnosis.*—

Uniform classifications of various aspects of a patient's behavior, disabilities, and social adjustments are increasingly needed in the mental health field to supplement data on psychiatric diagnosis. For example, classifications are needed to characterize the effects of mental disorders on activities of daily living, such as self-care, mobility, and communication; productive activities, such as work, homemaking and school attendance; and social activities, such as community and family life. Among the instruments that have been developed for assessing social adjustment are the Barrabee and Finesinger Social Adjustment Scale (19) and the Katz Behavior Inventories (20). The APA Manual provided a scheme for characterizing degrees of psychiatric impairment (4). Each of these instruments or scales has been used in somewhat limited situations. From the point of view of providing comparable statistical data that will permit assessment of the effectiveness of programs to reduce disability associated with mental disorders, it is essential to develop classification schemes that would

gain widespread acceptance and lead to the production of uniform and comparable data in this area.

8.0 Concluding Remarks

Within the past 3 years, significant national legislation has been enacted as part of an intensified attack on the problem of mental disorders. This has included the Community Mental Health Centers Act of 1963 (Public Law 88-164) (21) and the Community Mental Health Centers Act Amendments of 1965 (22). These two bills have provided the authority and funds for the development and staffing of comprehensive community mental health center programs for the provision of a wide range of coordinated mental health services in the community. Considerable data on the distribution of mental disorders and existing patterns of utilization of psychiatric services by various diagnostic groups of patients in the various age, sex, and socioeconomic groups of our population are required in the development and modification of State plans for these centers. Considerably more facts will be needed to measure the impact of these programs on changing the patterns of utilization of psychiatric services for various types of patients; assessing the effects on the patients, their families and the community, and the degree to which they may reduce the level of disability associated with specific types of mental disorders among the inhabitants of the communities so served. So that these facts will be meaningful and useful, it will become increasingly important to diagnose and classify patients accurately so that more precise data will become available on the effectiveness of various treatment modalities, therapeutic and rehabilitation programs for patients with specific characteristics. Unless something is done to improve the reliability and accuracy of diagnostic and classification methods, it will be impossible to determine

the changes specific for diagnoses with any degree of assurance.

The comprehensive community mental health programs are encouraging the development of psychiatric services in general hospitals. Already 1,043 of the 6,000 general hospitals of the United States admit psychiatric patients routinely for treatment and/or diagnosis and another 2,100 admit patients on an emergency basis only. The diagnoses on these patients are recorded in terms of the "Standard Nomenclature." To qualify for participation in Medicare benefits, psychiatric hospitals will have to be accredited by the Joint Committee on Accreditation. One of the recommended standards for accreditation is that the hospital should use the "Standard Nomenclature" as its diagnostic nomenclature. In addition, increasing numbers of commercial insurance carriers, as well as health insurance, group practice, and union health benefit plans are providing coverage for psychiatric disorders. All of this will result in greater usage of the existing psychiatric diagnostic terminology. Because of the difficulties in developing various kinds of case-finding techniques for general population surveys of mental disorders and because of the urgent need for data on psychiatric conditions for program planning and evaluation, and a variety of other purposes, it can be safely predicted that the diagnostic data derived from psychiatric facilities will be used increasingly for a variety of administrative, clinical, epidemiological, and research purposes. It is essential, therefore, that steps be taken to improve the ways in which the current psychiatric nomenclature is being used by the clinicians in our psychiatric facilities.

9.0 Summary

This paper considers the nomenclature and classification of mental disorders from the point of view of the biostatistician whose task it is to provide data on the incidence and prevalence of mental dis-

orders, patterns of utilization, and effectiveness of psychiatric services, and the fate of patients with such disorders in relation to time, space and demographic, socioeconomic, and related characteristics of patients and populations. The difficulties encountered in utilizing diagnostic data that originate in psychiatric facilities for developing national and international comparisons of diagnostic distributions of psychiatric disorders are discussed and illustrated. The critical problem in use of these diagnostic data stems from observer variability; i.e., the reliability and accuracy of the diagnostic terms recorded on clinical records and the consistency with which clinical psychiatrists use diagnostic terms and apply a given diagnostic term to patients with the same clinical characteristics or profiles. The need for developing diagnostic methods and instruments that would lead to greater precision and consistency in the way psychiatrists use diagnostic terms is discussed. It is suggested that studies of observer variability be carried out in which psychiatrists assess in a systematic way data obtained on sets of patients in a standard way. The data would consist of a standardized anamnesis, social adjustment history, and video tape of a clinical interview. Data derived from independent assessments of these materials by different sets of clinicians would make it possible to determine variation among clinicians in the use of diagnostic terms as well as in the profiles of patients on which there are varying degrees of diagnostic agreement among observers. Such research could lead to development of methods for obtaining greater objectivity and uniformity in classifying patients either on the basis of diagnosis or similarity of profiles. It will take time to develop ways for applying these methods to general use. In addition, the possibility exists that such diagnostic methods, when developed, could be utilized only in well-controlled research settings. Certain practical steps are, there-

fore, suggested on ways to obtain greater uniformity in psychiatric diagnostic practice and in utilization of diagnostic terminology through the use of diagnostic manuals and medical records committee at the hospital or facility level and at the community level. Some of the needs for classification of aspects of psychiatric morbidity other than diagnosis, such as types of disabilities and community adjustment, are also discussed.

REFERENCES

(1) Group for the Advancement of Psychiatry, Committee on Preventive Psychiatry. "Problems of Estimating Changes in Frequency of Mental Disorders" (Report No. 50): 469-521, New York, 1961.

(2) LIN, T., and STANDLEY, C. C.: "The Scope of Epidemiology in Psychiatry." Public Health Paper No. 16. Geneva, World Health Organization, 1962. 76 pages.

(3) PLUNKETT, R. J., and HAYDEN, A. B., editors. "Standard Nomenclature of Diseases and Operations" (fourth edition). New York, Blakiston, 1952. 1,034 pages.

(4) American Psychiatric Association, Committee on Nomenclature and Statistics. "Diagnostic and Statistical Manual: Mental Disorders." Washington, 1952. 130 pages.

(5) World Health Organization, "Manual of the International Statistical Classification of Diseases, Injuries, and Causes of Death" (seventh revision). Geneva, 1957. 393 pages.

(6) MORIYAMA, I. M.: "The Classification of Disease—A Fundamental Problem." *Journal of Chronic Diseases*, II (5): 462-470, 1960.

(7) U.S. Department of Health, Education, and Welfare, Public Health Service, National Center for Health Statistics. "International Classification of Diseases, Adapted for Indexing Hospital Records by Diseases and Operations." (PHS Publication No. 719), Washington, Superintendent of Documents, U.S. Government Printing Office, 1962.

(8) STENGEL, E.: "Classification of Mental Disorders." *Bulletin of the World Health Organization*, 21: 601-663, 1960.

(9) REID, D. D.: "Epidemiological Methods in the Study of Mental Disorders," Public Health Paper No. 2. Geneva, World Health Organization, 1960. 79 pages.

(10) BAHN, A. K., GORWITZ, K., KLEE, G. D., KRAMER, M. and TUERK, I.: "Services Received by Maryland Residents in Facilities Directed by a Psychiatrist." *Public Health Reports* 80 (5): 405-416, 1965.

(11) BABIGIAN, H. M., GARDNER, E. A., MILES, H. C., and ROMANO, J.: "Diagnostic Consistency and Change in a Followup Study of 1,215 Patients." *American Journal of Psychiatry*, 121 (9): 895-901, 1965.

(12) KRAMER, M.: "Some Problems for International Research Suggested by Observations on Differences in First Admission Rates to the Mental Hospitals of England and Wales and of the United States." *Proceedings of the Third World Congress of Psychiatry*. Montreal, Canada. 153-160, 1961.

(13) "The Registrar General's Statistical Review of England and Wales for the Year 1960." London, England, 1964, 191 pages.

(14) ROTH, M.: "Mental Health Problems of Aging and the Aged." *Bulletin, World Health Organization*, 21: 527-561, 1959.

(15) GREENHOUSE, S. W., and GEISSER, S.: "On Methods in the Analysis of Profile Data." *Psychometrika*, 24 (2): 95-112, 1959.

(16) SPITZER, R. L., FLEISS, J. L., BURDOCK, E. I., and HARDESTY, A. S.: "The Mental Status Schedule: Rationale, Reliability, and Validity." *Comprehensive Psychiatry*, 5 (6): 384-395, 1964.

(17) Joint Commission on Accreditation of Hospitals. "Standards for Hospital Accreditation." Chicago, 1960, 10 pages.

(18) Community Mental Health Centers Act of 1963, Title II, Public Law 88-164, Regulations. *Federal Register*, May 6, 1964.

(19) BARRABEE, P., BARRABEE, E., and FINESINGER, J. E.: "A Normative Social Adjustment Scale." *The American Journal of Psychiatry*, 112 (4): 252-259, 1955.

(20) KATZ, M. M., and LYERLY, S. B.: "Methods for Measuring Adjustment and Social Behavior in the Community: I. Rationale, Description, Discriminative Validity, and Scale Development." *Psychological Report*, 13: 503-535, 1963.

(21) Mental Retardation Facilities and Community Mental Health Centers Construction Act of 1963, Title I, Public Law 88-164, Construction of Research Centers and Facilities for the Mentally Retarded. 88th Cong., S. 1576, Oct. 31, 1963.

(22) Mental Retardation Facilities and Community Mental Health Centers Construction Act Amendments of 1965, Public Law 89-105, 89th Cong., H.R. 2985, Aug. 4, 1965.

The Role of Classification in the Development of the Science of Psychopathology with Particular Reference to Research¹

*David Shakow, Ph. D.*²

Introduction

D. Hack Tuke in "A Dictionary of Psychological Medicine" published in 1892, introduced his section on "Classification" with the following remark: "The wit of man has rarely been more exercised than in the attempt to classify the morbid mental phenomena covered by the term insanity. The result has been disappointing" (68). The present situation in psychiatric classification is, despite some advances, unfortunately not very different. The purpose of this conference is to examine proposals and projects, both past and present, which might result in reducing the persistent dissatisfaction with the categorization of mental disorders.

Although I am not a specialist in this field of classification, it would have been impossible for me, in the course of my involvement in psychopathology, not to have become closely enmeshed in diagnostic problems. This was true even before I went to Worcester, it was particularly true of my Worcester days and has continued, to some extent, in my post-Worcester period. The problem of achieving scientific rigor in the study of psychopathology has always been a professional concern of mine. I have, however, continued to indulge myself in the characteristic, but seemingly unavoidable, cycles of the experimentally oriented clinician: the alternation of overemphasis on meaning at the expense of rigor, and the overemphasis of rigor at the expense of meaning.

A revealing incident related to this problem occurred when I became a participant in Boring's seminar entitled, "On the Nature of Control in the Psychological Experiment" during the period of 1925-26. I was, to my delight, accepted for this seminar although I was only in my first year of graduate work. At the first session of the seminar, Boring introduced each of the members of the seminar to the group as a whole with a typical Boringian précis of the professional background of each person present. (Among the participants were Wever, Zener, McFarland, and several others who later made significant contributions to psychology.) When he presented me he said in essence: "Shakow is now in graduate work, living at the Psychopathic Hospital and interested in psychopathology. Why he is concerned with the problem of control I am not quite sure, but we welcome him anyway." This benevolently sceptical comment reflects the contemporary—and I fear even the present—attitude of academic psychologists toward the study of psychopathology and the work of clinicians in general. It illustrates their suspicion that anything that smacks of the clinical orientation is contaminated by looseness and verbosity as opposed to rigorous scientific control—the focus of Boring's seminar.

But to get back to the topic. Let me tell you from the very beginning what I plan to consider in the present paper. First I wish to say something about the general attitude I believe the investigator should take in confronting the problem of classification. This will be followed by my own personal associations with the problem, which I shall discuss in some detail. By "personal" I mean the attempts to deal with classification that have occurred in the contexts of the programs with which I have been involved. This will not only meet the assignment given me by the organizing committee of this conference, but, more important will present pertinent facts related to classification that have never been made public, and

¹ Presented in part as the Helen Sargent Memorial Award Lecture at the Menninger Foundation, Sept. 27, 1965.

² Laboratory of Psychology, National Institute of Mental Health, Public Health Service, U.S. Department of Health, Education, and Welfare.

which I think belong in the history of the problem. (A very brief résumé of more recent efforts in this area will follow.) I will then consider the various underlying issues of the classification problem as I see them. In my conclusions I will summarize some of the things I have learned during my association with the categorization of mental disorders: both the problems which are involved and the cautions which must be kept in mind when approaching this area.

General Attitude to Take

I do not understand how a psychologist can adopt a professional orientation which does not acknowledge classification as fundamental in dealing with the multiplicity of phenomena involved in the diagnosis of mental disorder. Classification is essential whether it is being made for therapeutic or for research purposes, but it is especially important for the latter. The objections that are directed against a particular classification system usually arise because the criticized form of classification differs from the one used by the critic himself. No matter how vehemently a psychiatrist or psychologist may oppose an existing system of classification, he almost always has his own method of categorizing patients.

Some persons connected with the field object to what they call the "premature" classification of psychiatric phenomena, a classification adopted before what they consider sufficient information had been accumulated to justify such an attempt. The decision as to what is premature, especially in such a subjective field as mental status evaluation, is based primarily on the personal opinion of the individual investigator. It seems to me that it would be a mistake for us to neglect any attempts at psychiatric classification, as long as they are based on a background knowledge of the phenomena involved. The systematic study and thought exerted in these past and present attempts at categorization have resulted in advances

in our knowledge of mental disorders and will, I believe, continue to provide valuable new insights in the field of psychopathology. It is required, however, that the proponent of a new method of classification present his system in clear fashion to the relevant professional community and permit its validity and usefulness to be evaluated by the ensuing scientific controversy. Such a disputation is characteristic of scientists' reactions to new methods and theories, and is, in the long run, frequently able to winnow out the wheat from the chaff.

Science is not possible without classification. It is essential to the objective investigation which is the core of the scientific method. The major objections to the classification of mental disorders which have been raised by many authors, among whom the most recent are Menninger and his group (43), are problems that almost inevitably result from any attempt at categorization. Such problems are reification—dealing with the abstract conception of the disorder rather than the actual behavior or symptoms presented by the individual patient; partialization—taking only a part for the total picture; privacy—the use of personal rather than public categories; and simplification—the substitution of a simple, easily comprehensible explanation for the complex of phenomena which are difficult to grasp in their intricacy. With the growing sophistication and critical self-evaluation of workers in the field, however, such obstacles will gradually be overcome.

Attempts at classification have been looked down on in other fields besides psychopathology. Among biologists, taxonomists were for a long time relegated to low status, as reflected in the appellation of "mere taxonomists" because of their narrow empirical approach. In recent years, however, they have introduced new principles of classification based less on the visually obvious structural criteria of previous systems, and more on the sophisticated functional-dynamic criteria from genetics, cytol-

ogy, ecology, and other branches of biology. This new orientation of classification is much more meaningful in the context of theory. One result of this updated classification process, which has been called the "New Systematics," has been the taxonomists' rise in status to the point where taxonomy has, according to Huxley, become "one of the focal points of biology" (23).³

The classification problem in psychopathology should, I believe, be approached in a similar fashion. The question before us is: How can we evolve a system which circumvents the kind of pitfalls we have mentioned, but still makes our work in therapeutics, and particularly in research, easier—in fact, possible? I hope that by the end of this discussion I will be able to suggest several principles which might serve to guide us in the development of such a classification system.⁴

At this point I cannot refrain from illustrating one of the objections I listed earlier by describing a method of classification which, although satisfying certain individual requirements of the person who devised and used it, does not meet the obvious criterion of public usability.

Many years ago I visited the apartment of a friend who had a large collection of novels and other books in his extensive library. Most of his volumes were the standard paperback editions of books published in foreign languages, which he had had bound in half leather of different colors. In addition to his professional psychological interests, my friend was an expert in the field of aesthetics and was, in addition to his other talents, a reasonably good painter. In organizing the books in his library, he departed from the traditional Dewey system of categorization and even shunned a simple alphabetical ordering. Instead, he arranged his books by the color of the

leather bindings, so that the volumes on his shelves made an exciting geometrical pattern of color masses—actually a bibliophilic Mondrian. Judged aesthetically, my friend's classification method was a great success. It was also a personally satisfactory arrangement because he knew his books intimately and could locate the one he wanted without difficulty. It is obvious, however, that such an idiosyncratic system of classification was not at all adequate for public use since it was designed solely to fulfill the particular requirements of a private person, rather than the needs of the public at large, indeed, of even any significant segments of that public.

Since the diagnostic aspect of the classification problem is so frequently emphasized and discussed, it is important to point out that diagnosis per se is only one stage of the classification process. There are actually three different stages of the process in psychopathology which call for our consideration: (1) The accurate description of the phenomena exhibited by the patient, (2) the syndromization of these collected descriptions and (3) the actual process of assigning patients to different categories. I shall have occasion to deal with each of these points in some detail as I develop my argument.

The Worcester State Hospital Experience

In discussing the problem of classification, I think I might do it most satisfactorily by using myself as a case study, even though it might be at the risk of revealing my narcissism. Since I have, in one way or another, been concerned with the diagnostic problem for a long time I feel justified in taking the risk!

Although my most intimate involvement with psychiatric classification came after 1928, even before that time I was confronted with some of its basic issues. It was during my first year in college, when I used to visit McLean Hospital quite regularly and was occasionally allowed to sit in on diagnostic

³ In the same context, Simpson's book (63) and article (64) are of relevance.

⁴ Mayr's chapter on "Behavior and Systematics" (41) has an interesting discussion of the problem of classification of behavior as related to systematics.

conferences, that I was first impressed with the difficulties in diagnostic procedures which were revealed at those sessions. This was especially true after the heady glamour of a lowly freshman's attendance at such exalted occasions had subsided. Later in my college career I attended a number of case conferences at the Boston Psychopathic Hospital where I observed similar problems. After receiving my bachelor's degree, I spent 15 months at the Worcester State Hospital during which time I regularly attended case conferences and "ward walks." I also carried out intensive personality studies, both structural and psychodynamic, of a few patients under the guidance of Dr. Grace Kent. These detailed studies particularly impressed me because of the discrepancies between my findings and the data presented in the regular conferences. When I returned to graduate work at Cambridge, I spent my first year living at the Boston Psychopathic Hospital. During this time I attended a number of the case conferences conducted by MacFie Campbell or, on occasion, by Karl Bowman. Although I was impressed with many of the aspects of these presentations I was still left with misgivings about the then current process of psychiatric classification.

My period of most intense involvement with the classification problem began in 1928 when I returned to the Worcester State Hospital. One of my responsibilities as chief psychologist was the direction of the psychological investigations conducted by the schizophrenia research service, a service which had been in existence for somewhat less than 2 years at the time of my arrival. Those first few years had been a kind of a "pseudopodic" period in which the approaches had been highly probative. The program had relied on the part-time work of the psychiatrists of the regular hospital staff aided by Anton Boisen—then the chaplain—and several of his associates.

At the time I joined the hospital staff, the research service was being expanded.

Professional personnel from various relevant specialties were being recruited, and in a short time an extensive multidisciplinary activity was in operation. This work lasted for about 20 years and was probably the most continuous and extensive research program on chronic schizophrenia ever undertaken. Although the different groups of investigators examined the same patients, some of the work was unidisciplinary, some of it multidisciplinary (either concurrent or interrelated), and some of it truly interdisciplinary. In the last category I include those studies in which independent variables, whether physiological, psychological, or psychiatric, were set up and in which patients were studied with regard to dependent variables in one or both of the other realms.

We were immediately confronted with a problem basic to the theme of this meeting—that of selection. Since our research dealt with schizophrenia, we, of course, needed to know what schizophrenia was! It soon became obvious that we could not always depend upon the routine diagnoses made by either the house service or by the Boston Psychopathic Hospital from which we received a considerable number of our patients.

It is important to realize the attitude toward psychiatric diagnosis at the time our experimental work on classification was being done at Worcester. This was already a period of great concern in psychiatry with diagnosis and there was widespread dissatisfaction with the traditional diagnostic methods. Various studies investigating the accuracy of the diagnoses made at the Boston Psychopathic Hospital compared with subsequent diagnoses made by the staffs of various state hospitals had been carried out by Southard & Stearns in 1914 (64), Lowrey in 1919 (36) and Wilson and Deming in 1927 (70). It was also at this time that Richard Cabot (7) published an extensive study comparing the original and autopsy diagnoses of 3,000 cases of physical disease.

Our examination of the psychiatric studies in this group made it plain why, for research purposes, we could not depend on the diagnoses made on our patients at the Boston Psychopathic Hospital. During the period of our own struggle with the classification problem, the article by E. B. Wilson and Julia Deming appeared in which the authors made a statistical comparison of the Boston Psychopathic Hospital psychiatric diagnoses with those of several Massachusetts state hospitals for the period of 1925 and 1926. They found a 34 percent disagreement between the Boston Psychopathic Hospital's diagnoses and those made at the state hospitals. The greatest increase in disagreement, when compared with the study done 10 years earlier by Lowrey, was in the diagnosis of dementia praecox or schizophrenia. Wilson and Deming pointed out, however, that their median agreement of approximately 60 percent closely resembled that reported by Richard Cabot in his paper exploring the diagnostic pitfalls identified in a study of 3,000 physical cases coming to autopsy.

We were especially impressed in the Wilson-Deming data with the differences between the classifications at the Boston State Hospital and those made at Worcester. Of the 57 cases diagnosed dementia praecox at the Boston Psychopathic Hospital, the Boston State Hospital called only 12 percent dementia praecox, whereas of the 72 diagnosed dementia praecox at Boston Psychopathic Hospital, at Worcester we called 37 percent dementia praecox. On the other hand, of the 134 cases diagnosed by the Boston Psychopathic Hospital as manic-depressive the Boston State Hospital placed over 28 percent in this category, whereas of the 20 sent to Worcester State Hospital so diagnosed, only 10 percent were called manic-depressive. The wide variation between these two diagnostic categories made at the two hospitals—1 to 3 and 3 to 1—were, in part, an outgrowth of the psychiatric orientations of their respective superintendents.

Dr. May of Boston State Hospital (as reflected in his paper, "The Dementia Praecox—Schizophrenia Problem" (39)) was strongly influenced by Kraepelin's theory of dementia praecox and therefore included in this category only those cases who had a poor prognosis. The attitude of Dr. Bryan at the Worcester State Hospital was, conversely, more progressive. He recognized, for instance, the catatonic subtype as a division of dementia praecox having a generally good prognosis. At the Boston State Hospital, under the influence of Dr. May, this type was, on the contrary, often labeled "manic-depressive." The professional attitudes of these men were reflected in the different number of patients at the two institutions falling into each of these groups.

The Wilson-Deming finding that the percent of agreement on psychiatric diagnoses was no worse than the degree of agreement Cabot found in his study of 3,000 cases of physical disease at first tended to make us more tolerant of the wide discrepancy we found in the classification of mental disorders. It was soon obvious, however, that we could not tolerate a 40 percent discrepancy in agreement in a research program—even if we could convince ourselves of the legitimacy of the comparison. It seemed rather unreasonable to us to credit state hospital diagnoses with the same status as the more objective biological criteria supplied by an autopsy. After much study of the statistical data of the period, we arrived at the conclusion that a strenuous effort had to be made to improve the quality of the data that served as the basis for the initial diagnostic work on individual patients. This policy was adopted as an integral part of our research.

First we agreed on certain ancillary criteria for the selection of patients to be used in our research. These were directed toward eliminating obvious irrelevant complicating factors. We felt justified in doing this since we were concerned with investigating the schizophrenic process it-

self rather than obtaining an accurate sample of existing schizophrenics. Patients were to be excluded from the study if: (1) They were over 50 years old; chiefly because of the liability of organic changes which frequently occur with advancing age; (2) there was coexisting physical disease revealed by standard diagnostic methods; (3) we were unable to obtain an adequate history; (4) there existed a marked language difficulty; (5) mental deficiency appeared to be associated with the schizophrenic condition; (6) the symptoms were predominantly those of psychoses other than schizophrenia; (7) there was recent alcoholic or other intoxication; and (8) they were female. This last criterion was not based on prejudice! Actually we recognized that we were depriving ourselves of the study of an equally important part of our sample and a probably more interesting group. The restriction was adopted only because it simplified a study involving endocrine factors which were under special investigation in the early years of our study. The endocrine complexity of women is so much greater than that of men, that it seemed expedient to exclude them from an already complex study. (Later, I might add, we did carry out a study of female patients to provide a control group for special purposes.)

But what is more important in the context of the present meeting was our criteria for inclusion. By what criteria did we establish that a patient was schizophrenic?

The problem of classification from the direct clinical approach appeared to us to break down into several stages that finally resulted in a diagnostic categorization: (1) The initial mental status examination; (2) the decision about criteria for placement in the schizophrenic category and, beyond that, in a subtype group; and (3) the final diagnostic process—both for the individual diagnostician and for the whole group of investigators.

We felt it important first to study earlier approaches to the careful description of

patients and evaluate the merits of these attempts. I shall briefly outline the history of such investigations done in the general period prior to the Worcester program, the defects we saw in the current process of classification, what we considered as possible improvements, and how we tried to achieve these improvements.

Previous to our own efforts, attempts had been made to develop descriptive rating scales to help with the classification problem. The works of Kempf (28); Plant (48); and the Boston Psychopathic Hospital study of Kasanin, Bowman and their associates (6, 26), were the only studies we found in which the attempt to describe psychotic patients had reached the stage of formal presentation.

E. J. Kempf, at Hopkins, constructed a pictorial behavior chart of mental diseases in the tradition of the work which Adolf Meyer had carried out at Phipps, following on work previously done at Kankakee, Worcester, and in the New York State Hospital system. Ward attendants were trained to record daily the presence or absence of certain traits in the patient's behavior that were listed on the chart. The list of 57 items was divided into two main types of behavior: spontaneous and required behavior. Such items were used as "destructive," "angry," "apprehensive," "reading," "entertainment," "spoon fed." Whether or not a standard for deciding on the presence or absence of these reactions was provided for the attendant is not known.

Plant's method was devised for use of the nursing staff at the McLean Hospital and closely approximated a degree rating scheme method. It consisted of 19 types of reactions with 10 to 12 subdivisions for each. The subdivisions consisted of statements describing concrete forms of behavior which involved little subjective interpretation on the part of the rater. The types of behavior covered were: attitude toward taking food, reaction toward nurses, emotional reaction, hallucinations, recent memory.

A rating scheme emphasizing personality traits was constructed at the Boston Psychopathic Hospital. Since it had been devised for a study of the sociological aspects of mental diseases (26), the scale emphasized this kind of material and the ratings were done primarily by social workers. The scheme was the most carefully constructed of the rating methods we investigated. Unfortunately, the ratings were based on historical material from more or less biased observers (relatives, friends, etc.), so that the prejudice represented went beyond the difficulty encountered by even an unbiased person in the attempt to report objectively in the complex subjective area of personality evaluation.

A familiarity with the then current mental status evaluation techniques routinely used by psychiatrists or even those employed in elaborately worked-up case studies could not but leave us dissatisfied. Our analysis of diagnostic methods uncovered several reasons for this dissatisfaction. First of all the methods lacked quantifying possibilities. The absence of a system for the quantitative determination of an individual's clinically derived psychological characteristics made it difficult to correlate these factors with quantitatively expressed variables such as the different aspects of experimental psychological and physiological functioning which were a major part of our research. The use of the traditional classification of types of psychoses was also, in general, unsatisfactory—a dilemma frequently recognized by the investigator and not infrequently accompanied by an apology.

There was also an inexactitude of definition in the qualitative use of the terms used for psychiatric symptoms and traits. From the psychological viewpoint the definitions used were often inadequate and not entirely reliable. Individual psychiatrists were found to characterize the same behavior differently in different patients or even at different times in the same patients, or different psychiatrists would describe

divergent types of behavior with identical or similar terms.

We believed that an adequate rating method would offer the most satisfactory quantitative technique for certain aspects of psychiatric work. Psychiatrists had, in general, shied away from any attempts at the quantification of their professional techniques and have long and vehemently argued against what they termed "oversimplification." While acknowledging the theoretical desirability of accuracy and quantification, they had tended to ignore these criteria as an impractical "counsel of perfection." There was, of course, some justification for their distrust of oversimplification and the reliance on quantitative rating. Nevertheless, it would be difficult for anyone involved with psychiatric classification not to agree with Gordon Allport's opinion that, "Notwithstanding the dangers and difficulties encountered in devising and employing rating scales, we are forced to recognize this method as the only available objective criterion of personality. The sources of error must be gradually overcome by the improvement of the technic of rating" (2, p. 449).

Some of the specific advantages of a workable mental status rating scale which particularly impressed us were the following:

1. The roughly quantitative material on psychiatric characteristics obtained from it would be suitable for correlation with data from quantitative studies in other areas made on the same patient. For a single patient, one could use the ratings of one psychiatrist, or the composite (average) ratings of a group of psychiatrists (which has been found to be more reliable from a technical standpoint).

2. Reasonably dependable qualitative material for following up the same cases would be made available.

3. Another advantage would be the partial elimination of the personal factor which interfered with the objectivity of

the then current evaluative system. Although this personal factor might be important in therapy it often proved to be a marked handicap in investigation.

4. A working mental status scale would offer more dependable material for the determination of syndromes, and even in the determination of diagnostic types.

5. It would also offer a method for evaluating psychiatrists. One of the primary qualities of a good psychiatrist is his ability to utilize his "peculiar" professional language accurately in the description of the patient. The reliability of this use of professional phraseology could be checked by his successive ratings on the same patient and the validity of this usage could be measured against the composite ratings on the same patient of a group of other psychiatrists of known ability.

6. It would offer as well a method for training psychiatrists.

(a) The use of a mental status rating scale emphasizes the need for accuracy in symptom description and classification.

(b) It would make it possible for an individual student to check his use of psychiatric terms against the ratings of experienced observers on the same patient.

(c) The mental status scale also keeps the psychiatrist alert to the particular items of information which he should try to obtain during the course of a mental status examination. Eventually, the scale might even serve as the basis for a standard interview that preserved the flexibility and open-endedness in administration that any reasonable use of the interview technique calls for.

With these goals in mind, we undertook several preliminary studies. For the first of the studies, we made up a list of 104 traits, symptoms, and response patterns (which I shall hereafter call traits) of all types of psychoses. These were based on data taken from leading psychiatric texts, and the contemporary periodical literature. Kraepelin, Bleuler, W. A. White, Jelliffe and White, Kempf, Rosanoff, Henderson

and Gillespie, Strecker and Ebaugh were among the authorities consulted. The clinical experience of the psychiatric and psychological staff at the hospital was also called upon in the construction of the initial scale. First attempts at definitions of these traits were adopted from textbooks, dictionaries (psychiatric and psychological), and other similar sources and a scale of the degree of the presence or absence of the different traits was established. In general we tried to fit the traits into a five-point scale with descriptive characterizations for the different degrees wherever possible. (There was also a 0 category to indicate that a certain trait could not be rated.) In order to eliminate any possibility of the rater consistently checking one of the extremes, the "most marked" presence of a trait was sometimes set at 1, sometimes at 5. Rating forms for the traits and instructions for the use of the scale were gradually developed, together with a trait dictionary containing the definitions of both the traits and their degrees of intensity. Approximately 70 patients were rated by the scale once, some 25 of these taking part in two studies spaced 3 months apart. The comments and criticisms of the psychiatrists and others using the scale were collected for each of the traits. A preliminary statistical analysis was made of the collected ratings to determine such problems as the items which were difficult to rate and the items on which there was considerable disagreement. On the basis of our work with the first preliminary form, we found that the definitions of the traits were the most unsatisfactory part of the scale. Modifications of this aspect of the rating device were, therefore, given primary attention in the construction of the second scale.

When the material was collected and the comments and criticisms organized for the second study, the four psychiatrists, the chief biometrician, the resident director of research, and I held an extended series of active conferences on the project. As a

result of these discussions, a second rating scale was organized along the following lines.

1. After detailed discussion on each characteristic, the definitions were changed to meet the consensus of opinion of the group. Considerable weight was, however, accorded by the group to the opinion of authorities in the field.

2. Another result of the discussions, was the addition of many new items and the omission of a few, the final number totaling 138.

3. In this scale the descriptive characterization of the degree of presence of traits, except for the normal distribution basis represented by the categories of "very marked," "marked," "considerable," "some," "none" was eliminated. It was felt that this part of the scale should depend on a large accumulated body of clinical experience and, that, while this was being collected, it was best to leave the matter open.

4. To obtain this type of information, each staff member working on the second study used a notebook in which he recorded comments, difficulties, or criticisms that occurred to him while using the scale. We hoped in this way to compile a detailed description of each degree of a single psychiatric trait, together with a series of actual examples of behavior included under each degree. The primary purpose of making such detailed characterizations of the traits was to reduce the element of personal bias in determining the intensity of the traits—probably the most difficult part in the construction of any personality rating method.

5. New forms, a new set of instructions, and a new dictionary were organized. A special place was left in the new form for the rater to record sources of information on which the judgment was based. We hoped that after a time it would be possible to specify the sources of information for most of the traits, and that we might even weight the ratings according to source.

6. The scale was subsequently used in our research program. Because of other involvements, however, except for minor modifications, the plan to refine the system further was never executed. (A separate well-designed project was formulated in an attempt to develop the rating scheme further, but we were unable to put this into effect because of lack of the needed outside financial support.)

In the course of work on this project, we recognized various sources of error common in such an enterprise and the need for certain controls to meet them.

1. Differences in the length and depth of acquaintance of the psychiatrists with their patients was one source of difficulty. We partially controlled for this through corrections based on a statistical study of the effect on the ratings of the varying lengths of time the psychiatrist had a particular patient under treatment.

2. The five-point "very marked" to "none" scale to describe to what degree traits are present appeared inadequate. A more satisfactory detailed descriptive scale for each of the degrees based on examples given in textbooks, personal experience, etc., was continuously being evolved.

3. Differences in the difficulty of rating traits were another source of error. Traits were defined and redefined, but if it was impossible to achieve agreement, the given trait was finally eliminated.

4. To counteract the "halo" effect in rating, traits were defined as exactly as possible. The insidiousness of this personal effect on rating was stressed in training the observers.

5. Another source of error was the observer's use of "expected" rather than actual performance as the criterion for rating. Emphasis was repeatedly placed on the importance of making actual performance the basis for a rating in order to determine the bases on which judgments of behavior were made. The psychiatrists were frequently called upon to give the rationale for their particular rating of a trait.

6. Variations existed in the standards used by individual psychiatrists. There was an attempt to correct for these differences by making distribution curves of each observer's ratings for each trait rated. The need for constancy of standards was stressed in the training of the observers. Rechecks were made of previous ratings at regular intervals and emphasis was placed on the use of the average as the standard.

7. Ratings were always made independently to prevent the influence of raters on one another.

8. To control the influence of previous results on successive ratings, the raters were instructed to fill out the forms without referring to earlier ones they had filled out on the same patient.

9. The intraindividual variability of each patient also posed a constant problem. The instructions to the raters were, therefore, to consider variability as an expected characteristic, but one calling for continuous scrutiny.

10. Frequent conferences on procedure and additional training sessions were held to detect general errors which tended to creep into the work.

The material was then treated statistically. The following points which resulted from this statistical analysis were among those kept in mind by the participants in the project:

1. Which items lent themselves to a rating scale technic?

(a) Items which consistently yielded zero ratings (i.e. "cannot be determined"), even after attempts at redefinition, were eliminated.

(b) The differences among psychiatrists might occasionally indicate that one or another had a superior technique for dealing with certain items. A subsequent discussion of these differences often resulted in a redefinition of the item.

2. The reliability of ratings took into consideration the variation of the individual psychiatrist, the variation among

psychiatrists, and the study of differences in successive examinations by the same psychiatrist. These variations were carefully analyzed to make certain that they were not due to shifts in scoring standards.

3. Personal bias was dealt with by self-ratings of observers which were compared with ratings by other psychiatrists and sophisticated raters not engaged in the psychiatric work.

4. The statistical determination of the syndromes used such methods as inter-correlation of the various traits. The traits resulting from this statistical treatment were again redefined and their degrees more descriptively stated. There was also some attempt to weight traits by the pooled judgment of experts on their relative importance, or their reliability, and by a combination of both.

It should be noted that thus far we have been discussing only the first of the three aspects of the problem of classification—the raw data of the characteristics which make up the syndromes and how they might be obtained with some reasonable degree of dependability. We can now go on to the next step of development of the diagnostic process: how we determined the syndromes of traits which were to be the basis for diagnosing a patient as schizophrenic and placing him, whenever possible, in one of the subcategories of schizophrenia.

Here too we relied on the standard texts and other literature, together with the experiences of our own staff. We compiled a list of symptoms which were regarded as characteristic of the schizophrenic disorder and additional lists for the various standard subtypes of schizophrenia: simple, catatonic, paranoid, and hebephrenic. We circulated this diagnostic scheme among the psychiatrists and other members of the staff and finally compiled a list which seemed generally acceptable and which all participants in the study agreed to use as the criteria for diagnosis.

Now, having both the raw traits for characterization and the basis for combining these into syndromes, we came to the third aspect of the diagnostic process—the actual procedure we utilized in the final classification of the patient. The procedure I shall describe was arrived at only after considerable trial-and-error experimentation with other methods. The classification of a patient was decided upon at a diagnostic staff conference which was attended by the four psychiatrists, the director and the associate director of research, the chief of internal medicine (the last three being internists), and the two senior psychologists. Information on a particular patient was presented by one of the four psychiatrists to whom that patient had been assigned in accordance with the established procedure of rotation. He discussed the psychiatric material on the patient together with the data from the patient's social history. This was followed by a presentation of the results from an extensive battery of psychological tests. The patient was then interviewed before the group (unless contraindicated by his condition), and an interpretative discussion of the case followed.

On the basis of these data, three decisions were made—about diagnosis, prognosis, and suitability for the research program. All these decisions were carried out anonymously. I wish strongly to emphasize this point with regard to anonymity, for its value should be clear to those who have had experience with the manner in which staff conferences are ordinarily conducted, and the errors introduced under these conditions. In his personal evaluation, each member of the staff wrote down his opinion of the patient under consideration on the three topics mentioned and passed it on to the chairman of the meeting. The final results of the vote were tabulated and announced by the chairman. There had to be a practically unanimous agreement about the suitability of the patient for the research and about the diagnosis of schizo-

phrenia. With regard to subdivision, the standard subtype categories were used if there was substantial agreement on the subtype classification of the patient. If not, then the categories of "mixed" and "unclassified" were used liberally. In addition, we developed a category called "late indeterminate." This subdivision was used for chronic patients who had at one time clearly belonged to one of the standard subtype classifications but who did not currently demonstrate any of the classic symptoms of that subtype.

If accepted as suitable for the study, the patient was interviewed daily by the psychiatrist throughout 3 months of active investigation, and notations, which were incorporated into weekly notes, were made of his mental state and behavior. The rating scale system previously described was used over a considerable period of the study. When necessary, changes in the original diagnosis and subtype category of any patient were made at diagnostic conferences assembled for the purpose. In addition to the psychiatrist's observations, daily notes were also kept by the nursing staff and by two specially trained ward observers who were instructed to report selected aspects of a patient's behavior. Notations were made less frequently during the rest periods—the periods in between those in which active study of the patient was going on.

And what did this effort in classification amount to—what results did we achieve? It is my opinion that, in spite of its incompleteness and its many failings and inadequacies, the study was one of the most carefully carried-out, large-scale projects ever attempted in the classification of chronic schizophrenia. I must point out here that since our study was directed at the chronic group, our problem was made much easier than if we had been dealing with the whole range of mental disorders, or even with only the full group of schizophrenic or schizophreniform conditions.

The results of the project may be considered under the headings of both overall and specific effects. In general the atmosphere of the study fostered intimate knowledge of the individual patients, accuracy in clinical investigation, and it called for objective evidence in establishing the criteria for diagnosis as well as emphasizing the importance of unambiguous psychiatric categories. The research group developed a sensitivity both for the complexity of the problem and for the need for accurate characterization of the syndromes.⁵ The project resulted in the evolution of a common area of agreement about psychiatric classification among the psychiatrists, as well as in the rest of the staff, despite the different philosophic backgrounds represented, varying from a psychoanalytic approach through a structural to an organic orientation. In short, we were able to minimize significantly what is inevitably a sort of Gresham's law of psychiatric discourse—the tendency for loose talk to drive out rigorous talk.

The specific effects of the project were many. First, criteria were developed which offered considerable reliability in evaluation. We were also able, on the basis of these evaluations and classifications, to establish reasonably homogeneous groups of psychiatric subtypes so that these subtypes correlated with quite independent psychological studies made of these same patients. We were able to differentiate the schizophrenic group both from other psychoses and from normal subjects, and even differentiate the various subtypes within the psychosis.

During an experiment using the play technique (61), Rosenzweig and I observed how clearly normal, paranoid, and hebephrenic subjects differentiated themselves from each other on the basis of their be-

havior in the assigned task. Having selected the subjects for our experiment according to their mental status and the other criteria used in the psychiatric classification, we were impressed with how distinctly the independent psychological data differentiated these categorized groups. Another specific result we achieved was a fairly high correlation between items on the mental status examination and the results of a wide variety of psychological experiments and tests.

Needless to say, the study was terminated before completion. Such an outcome is almost inevitable with any project in this most complicated area. I am certain that even if we had obtained the necessary financial support for the elaborate study of the current mental status system I have mentioned, we would not have been able to achieve conclusive results. The task is an immensely difficult one, it has many steps, it takes a long time, and for that reason the battlefield is strewn with bodies of only partially viable projects abandoned to eventual death. I am convinced that only from a long succession of such systematically organized careful studies will any meaningful solution of this problem be achieved.

A question which probably occurs to many of you is why we did not use factor analysis in our study of classification. It seems so natural an approach, especially to psychologists. At the time we were planning the study, I discussed the problem with Morton Jellinek, who was chief of biometrics in our research group during eight of its most intensive years. He advised against the use of factor analysis. His reasons for discouraging us were similar to those that Thurstone gave in his own early discussions of the use of this technique, namely, that the results obtained from a factor analysis depended entirely on the quality of the data put into the statistical hopper (67). Jellinek agreed with us that our primary task at the time was to collect

⁵ Anna Freud on a recent visit to the National Institutes of Health in 1965 described what took place at Hampstead as a result of the establishment of their index—a marked increase in the understanding of the phenomena they were studying. This same general effect of our procedure was noticed in the Worcester group.

a body of dependable and accurate descriptive data.

This reference to Jellinek brings me to a consideration of some of the points about psychiatric classification made by him in a paper published in 1939 (25).⁶ I consider Jellinek's paper one of the important products of our program and of unusual significance for the field of classification. I recommend the entire paper to you as worthwhile. Since this work is little known and only rarely mentioned, I shall take the liberty of paraphrasing or quoting some of the major points.

Recently advanced arguments for abandoning classification in psychiatry are invalid. The classificatory system, however, must be placed on more secure foundations than in the past, the guiding principle of this system must be relevance. The criterion of relevance cannot be statistical * * *. Material knowledge, intuition and logical process alone can lead to the recognition of relevance; the hypotheses may then be validated statistically. Psychiatry may gain by greater attention to classificatory principles of other sciences, but it requires greater independence in the execution of these principles * * *. There is little doubt that a reform of classification is more likely than its entire abolition * * *

The argument of individual differences cannot be admitted as invalidating the purposes of classification. The variation of mental and emotional manifestations from individual to individual, though great, is not greater than individual differences in biochemical and physiological behavior * * *. As great and important as individual differences may be, certain common principles can be discovered which make generalizations possible and classification imperative. The most important consequence of the recognition of individual differ-

ences is that generalizations, whether qualitative or quantitative, should not be made from group studies but mainly from the joint experience of intensive individual studies. This is not less true of physiology than of psychiatry * * *.

* * * * *

Two needs are satisfied by classification—a minor and a major. The minor need is the convenient and easy handling of data * * *. Such classifications amount to glorified filing systems. The data become easily recognizable by some brief label which ordinarily pertains to some superficial, but rather obtrusive characteristic * * *.

The major need is to bring the data into a system which emphasizes either homologous or analogous characteristics as those which reasonably reflect common origin * * *. Actually, however, relevant classification is the result of deep understanding, and the source of even deeper understanding. Relevant classification will result as a minimum in an economy of thought. Relevant classification of the data is practically indispensable for the development of theories, although relevant classification itself reflects a theory * * *.

Jellinek then goes on to discuss the evolution of botanical taxonomy from superficial characteristics as the basis of categorization to natural and structural syndromization of related characteristics.

* * * * *

That syndrome should be an expression of common origin is a highly important idea. It is an idea which should be more stressed in psychiatric classification. A syndrome which is rooted in common origin complies with the most important requirement of scientific classification, namely, it has predictive value * * *.

* * * * *

Botany and zoology have stressed the building up of classification on true structural elements. In psychiatry, it would seem that an insufficient effort has been made to distinguish structural elements of personality from its arabesques. I believe that many of the

⁶ I must here pause to pay tribute to Morton Jellinek, whose death a few years ago was such a great loss to mental health. Before going on to become the world's leading authority on alcoholism, he contributed immeasurably to our research from his amazing store of wisdom as a biometrician as well as from his extraordinarily wide range of competence in so many other areas.

habits of patients used in psychiatric classification are nothing else but arabesques * * *. There was a time when birds used to be classified as climbers, runners, swimmers, etc. This classification was abandoned as irrelevant when it was realized that some birds of entirely different habits are much closer related than some birds of the same habits * * *.

There are obvious traits which may be relevant, but the most relevant we are likely to find at deeper layers only. This hidden likeness is a thing that is not manifest in a recognizable trait, but can be arrived at through abstractions only. These abstractions will take on the form of frank figures of speech and occasionally of analogies * * *. Analogies or figures of speech are the equivalents of pictorial symbols. Science operates with symbols but requires that these symbols should have an operational value * * *. When abstractions in the form of figure of speech are to be used in a classificatory system, it becomes imperative to establish their operational nature in contrast to their purely aesthetic or economic values * * *.

I do not say that psychiatry has not recognized the principle of relevance, but I contend that it has gone about it in a way which may be adequate for some sciences but not for psychiatry. The determination of relevance in psychiatry has been largely statistical. * * * There are many psychiatrists whose attitude is statistical. Their judgments are based on the fact that many, most, or all of their patients showed this or that characteristic or that "this trait was more frequently observed simultaneously with that trait than with any other one." This is what I mean by the statistical attitude. I doubt that relevance in psychiatry can be determined simply by the criterion of frequency or of association. The latter per se particularly does not establish relevance. An investigation would show that pencils of yellow wood finish have a significantly greater incidence than pencils of any other kind of finish. But is the yellow color essential to pencils, to writing, or to the communication of thought? The statistical determination of association may be a

supplementary test but not a proof of relevance. Psychiatry deals with abstractions, and these can be arrived at through reflection but not through counting. Ultimately, statistical methods may be invoked for the testing of hypotheses built by an intellectual effort * * *.

Psychiatric classification will gain much through giving closer attention to the broad principles of other sciences such as botany, but in the execution of these principles psychiatry must be more independent and original than it has been hitherto.

I would also like to cite certain relevant passages from an earlier paper of Jellinek's on a related topic, "Some Uses and Abuses of Statistical Methods in Psychiatry" (24), which he delivered at a round-table discussion of the 1937 meeting of the American Psychiatric Association.

Statistics is a system for testing of hypotheses but not a substitute for thinking * * *. In the determination of syndromes much more theoretical penetration than statistics offer is necessary * * *. There is before us the question of what elements or characteristics should enter into our syndromes and of how we should go about finding these characteristics which later we wish to bring into a statistical system. If in the end we must rely on some correlational measure then we must make sure from the start that such correlations as we obtain should represent syndromes of operational value * * *. I would regard this type of analysis as essential before the material is handed to the statistician, for his advice will always be based on the assumption that the data are meaningful * * *. Once the epistemological task has been accomplished, statistics can proceed to test hypotheses made and may even point in the direction toward new hypotheses * * *.

Have not too many of us, at one time or another, employed statistical analysis as a shortcut to avoid rigorous observational and theoretical work, or employed the magic of a statistical approach as a substitute for dealing with the epistemology?

And is this not one of the weaknesses of much of the current work with factor analysis? The relevance of this point to the others referred to in my excerpts from Jellinek's paper does not need further emphasis.

It was during this same period that Saul Rosenzweig, then a member of the Worcester group, published a paper on "The Basis for the Improvement of Personality Tests With Special Reference to the M-F Battery" (51). In this paper on the use of a test approach, Rosenzweig also affirmed the need for theory, rather than reliance on the pure empiricism which he noted as being employed by Terman and Miles in the development of their battery for the study of masculinity-femininity.

I must admit that Jellinek's insistence on the importance of theoretical generalization about the phenomena observed was not sufficiently stressed during our Worcester project. It was, however, implicitly emphasized in the context of the staff conferences in which we considered the symptoms of the patients in terms of such theoretical concepts as regression and restitution, rather than in their purely empirical aspects. Later, however, during our theoretical analyses of the results of our physiological (20) and, particularly, psychological studies, generalizing principles were made more explicit (55, 58, 59). It was also brought out in a preliminary unpublished effort to analyze modes of adaptation with special reference to defence mechanisms (57).

Despite efforts such as those attempted at Worcester, studies which demand high standards and care for successful execution, I confess that I have frequently been left with an underlying sense of uneasiness and ambivalence about the process followed. I have never been quite sure whether we were actually approaching the veridical through a process of mutual education in which the staff worked through the data and methodology to a generally acceptable system, or what we were engaged in was

rather a form of indoctrination in which the observers were both consciously and preconsciously constrained into a form of nominalism—agreeing to call something by a certain name whether accurate or not.

This faced us with what is probably the unsolvable dilemma of classification. I do not know what else we could have done. The two choices were either to approach the classification of patients from each participant's initial idiosyncratic frame of reference, which would result in a confusion of definition, or for all participants to accept either the discussion leader's, or the total group's, notion of the psychological phenomena or dynamics involved. These choices, perhaps, illustrate the fundamental paradox in education. The result of a truly educational experience is the synthesis of the student's uniquely individual contribution with the theoretical give-and-take of the group process. There is inevitably some loss of independence in group instruction which, however, is usually more than compensated for by the student's gain in knowledge and his development of the capacity for constructive critical self-appraisal.

The stages of the particular educational process involved in the development of a classification system seem to be essentially these: (1) The initial great differences in the criteria for diagnostic judgment which exist among the participants; (2) the minimization of these differences as a result of the participants' discussions of their differences in professional opinion, and the resulting achievement of an optimal degree of agreement and reconciliation; (3) then a dilemma arises: Are the judgments veracious, or are they merely group-accepted stereotypes? In the end, I suppose, we must accept a minimal level of stereotypy. This will not, I believe, be a critical disadvantage if the participants have already employed both self-criticism and mutual criticism in the attempt to clarify the concepts they use. Some of the obstacles in dealing with this problem lie

in the initial wide divergence in the professional knowledge and personal background and orientation of the participants in the project.

I trust you will forgive me for having elaborated at such length on this small segment of the history of classification. It grows out of the intimate knowledge I have of this program. If the discussion impresses you with the inordinate number of steps involved and the tremendous amount of work which such an enterprise demands, I will feel justified for having burdened you with this lengthy presentation of the Worcester program, a program which only partially achieved its goals. I believe, however, that what we learned from this program has relevance not only for the points I have already made, but also lays the ground for additional ones I shall present shortly, after briefly considering more recent developments.

Specific Efforts at Classification

In this section on more recent efforts at classification, I will touch only briefly on this complicated and extensive subject, for I am sure it will be dealt with in detail by other members of this conference. I shall, however (as a token effort toward completeness in my assignment for this conference), list some of the significant studies with which I am acquainted. As I have already stated, the process of classification involves three distinct stages which I shall consider separately.

The studies of Kempf, Plant, and the Boston Psychopathic group, three of the earliest efforts of recent times to describe individual characteristics objectively, have already been described. There have since been numerous other studies in this area. Some of the outstanding contributions to this aspect of the development of classification have been the Zubin Scale proposed in 1934 (73), the Malamud-Sands Scale (37) also developed at the Worcester State Hospital, the Monroe Scale (45), the Men-

ninger Scale worked out for their psychotherapy project (69), the scale developed by Phillips and Rabinovitch (49), Hamilton's revision of an early version of the Lorr Scale for use with chronic schizophrenics (18), the Lorr Scale (34), the Wittenborn Scale (71), the Rockland-Pollin Scale (50), and the Spitzer Scale (65). There are, of course, many other significant studies I have, of necessity, neglected in this brief list. Among the personnel used as raters in these scales have been nurses, attendants, ward observers, psychologists, and psychiatrists.

We find several kinds of approaches used in the study of the synthesis of personality traits into diagnostic categories. The oldest and most important is, of course, the combination into syndromes of clinically observed traits based on the day-to-day experience of the psychiatrists dealing with patients. This has, of course, been a continuing method almost from the infancy of psychiatry. There has also been a trend in recent years to use statistical approaches to classification in the form of factor analytic or cluster analytic techniques. The first effort of this kind was made by Moore (46) which has been followed by a succession of studies: by Degan using the Moore material (12), Guertin (17), Eysenck (14), Lorr (35), Wittenborn (72), and the related effort of Meldman (43).

The proposed systems of actual categorization break down into two major types according to the kind of data employed: the clinical material mainly provided by psychiatrists, and the statistical data based on ratings provided, with few exceptions, by psychologists. Thus, on the one hand, we have the present standard classification system used by the American Psychiatric Association, and, on the other, the proposed typologies derived mainly from factorial techniques.

I shall make no effort at special consideration of any of these attempts. Instead I shall refer you to some authors who have

already made such appraisals of one or the other of the two kinds of typologies: Marzolf (38), Stainbrook (66), Meehl (42), Eysenck (15), Lorr (34), and Overall and Gorham (47). The following papers deal specifically with the problem of factorial methods: Bordin (5), Humphreys (21), and Coan (9). Although I do not agree with all these authors have said, there is a wealth of information as well as an unusual array of perceptive evaluations in their discussions.

In making my own summary evaluation of the problem of the classification of mental disorders, I prefer to state my reactions to personal experiences at Worcester and elsewhere, and the development of my thinking on the problem.

Discussion of Problems of Classification

I will now consider some of the principles underlying efforts to classify mental disorders. First, I shall examine two basic philosophical questions involved in attempts at categorization and then go on to consider several specific methodological problems. Since I have found the statement of these problems in the form of dichotomies most helpful, I shall so consider them in the following discussion.

Let us dispose of a disturbing, though irrelevant issue, which is, unfortunately, often a factor in the use of classification. This problem, dealt with by Hunt, Wittson and Hunt (22), is the effect of classification employed for purposes of the practical disposition of institutionalized patients rather than in the effort to facilitate the achievement of fundamental therapeutic or scientific goals. Certain institutions pigeon-hole patients in categories which are not truly justified by their symptomatology to enable them either to hospitalize the patient without question, treat him as an outpatient, or dismiss him. Such arbitrary disposition of patients circumvents the true issues involved in classification. The intent

of this categorization may, of course, be humanitarian, such as a clinician's desire to avoid labeling a patient schizophrenic because of the possible repercussions of such a diagnosis in his future life. It is obvious, nevertheless, that such practices in classification are extraneous to our efforts to arrive at a satisfactory categorization of mental disorders and, therefore, do not belong in the context of this conference.

Two basic difficulties have particularly plagued the field of classification—the tendency to adopt extreme attitudes, and uncertainty about the model to follow. In earlier years there was, perhaps, too strong a trend to accept passively the classification of mental disorders proposed by authoritative figures in psychopathology without either sufficient questioning or the utilization of empirical methods to evaluate their validity. In more recent years, however, a tendency to adopt uncompromising extreme positions in relation to classification has developed. Such extreme trends in the effort to classify mental disorders are reminiscent of what Franz Alexander (1) has labeled “dialecticism,” and have resulted in either the acceptance of rigid categorization, or a nihilistic attitude expressed in the desire to discard all classificatory systems. The latter appears, at least in part, to be a consequence of the contemporary emphasis on psychoanalysis, and the related stress on the idiographic approach to personality study which has characterized this same period. The combination of the two has resulted in a scepticism about classificatory systems. The consequence has been a heated controversy over the value of the standard categories which is reflected in the conflicting attitudes of a Mayer-Gross (40) and of a Karl Menninger (44).

The other basic problem—considerably the more significant of the two—revolves around the question of which theoretical model should be employed in attacking the problem of classification. We touch here

on a basic issue for psychiatry. For many reasons, the major ones being historical-vocational, psychiatry has naturally identified itself with the medical disease model in which psychiatrists are still trained. This is true although the actual practice of psychiatry has, from quite early days (compare the period of moral treatment of the early 19th century), followed a psychological-behavioral model. The resulting dissociation has plagued psychiatry almost from its beginnings and has resulted in a split (which I have elsewhere used to describe schizophrenia) like that of the 19th century which, according to Santayana, "yearned with Rousseau or speculated with Kant, while it moved with Darwin, Bismarck, and Nietzsche * * *" (53). The problem which has recently been intensified by both the important position that psychoanalysis has achieved in psychiatry and the more recent development of social psychiatry, has now reached an incandescent state. Eissler's recent book, "Medical Orthodoxy and the Future of Psychoanalysis" (13) provides an elaborate statement of the effects of the split on psychoanalysis.

For this audience it is unnecessary to emphasize the great overlap of the behavioral and the medical in psychiatry and the profession's obligation to incorporate both aspects. The division which is made between these two aspects of mental disturbance is frequently artificial. Psychiatry stands at the interface between the biological and the psychosocial, and straddles both areas. This hazardous stance is especially uncomfortable for a field that is so overburdened with unsolved problems. Partially because this basic schism has not been resolved, psychiatry has been relegated to an altogether too low status both in its relationship to the medical sciences and to the basic behavioral sciences. (In passing, it should be pointed out that it has also resulted in a crisis central to the development of psychoanalysis—an issue with which the book of Eissler deals, and a topic

on which a number of persons, including myself, have spoken out.)⁷

Isn't it time for psychiatry to take a firm position on its relationship to both the biological and the psychosocial sciences, particularly the latter about which I feel competent to comment. This would require a reevaluation of training programs as well as psychiatrists' recognition of the importance of psychology and the social sciences to their activity. Although I do not insist on the narrow definition of a psychosocial scientist as one who has received an advanced degree in the discipline of psychology and the social sciences, I do, however, firmly believe that if a person is to engage in psychosocial work he must, whatever his original background, have substantial experience and training in the psychosocial field.

Many of psychiatry's difficulties have, I believe, arisen from the futile attempt of psychiatrists to stand astride both areas of the biological and psychological with no firm footing in either. The present situation makes me most sympathetic toward a training program like the one proposed by Kubie (32) for the adequate training of psychotherapists. Modifications would, of course, have to be made in Kubie's proposed curriculum if we were to construct a similar program suitable for the training of psychiatrists, but it might very well omit certain parts of the present curriculum and include much more relevant aspects of the biological, behavioral, and social sciences.

A fundamental social question which I have considered elsewhere (56) arises in this context. What is the justification for training psychiatrists in the standard medical curriculum when their practice is, with relatively few exceptions, and certainly for the foreseeable future will be, primarily in the

⁷ Shakow, D. (1953), "Discussion of Talcott Parsons' paper, *Psychoanalysis and Social Science*." In "Twenty Years of Psychoanalysis," editors: F. Alexander and Helen Ross. New York: Norton, pp. 216-226.

Shakow, D. (1962), "Psychoanalytic Education of Behavioral and Social Scientists for Research." In *Science and Psychoanalysis*, Vol. 5, editor: J. H. Masserman. New York: Grune & Stratton, pp. 146-161.

fields of the psychological and social sciences?

It is true that psychiatry historically has actually dealt with three areas of problems: (1) Those which are fundamentally physical in origin but which have psychological manifestations—the group of disorders which may be called somatopsychic; (2) those where the etiology is fundamentally psychological but which have physical manifestations—the group of disorders which are called psychosomatic; and (3) those which appear to be fundamentally psychological in nature and behavioral in expression—the group of functional disorders or maladaptions. The last group would appear to be the major sphere of psychiatry, while the first group of somatic disorders are frequently studied and treated by neurologists and other physicians. This has become increasingly true since medicine in general has developed a more sociopsychological orientation. Although psychiatry continues its involvement with the psychosomatic disorders, it is becoming even more concerned with personality disturbances resulting from psychosocial conditions.

Psychiatry must, it seems to me, straightforwardly adopt a social orientation to replace its reliance on medical tradition and the accompanying prestige resulting from association with the honored profession of physical medicine. In order to achieve this new position, psychiatry must reexamine the adequacy of the present education of its practitioners and must also deal with the conflicts arising from simultaneous identifications with the psychosocial disciplines because of the relevance of their activities and with the medical profession because of its specialized training—a training often remote from the sphere of personality disorders for which psychiatry has an important social responsibility.

I find it most difficult to agree with the arguments of persons like M. R. Kaufman (27) about the particular importance of the medical tradition. Psychiatry must, for the

sake of its own future development, clarify its relationship to the behavioral and social sciences particularly, and reorganize its training programs with this end in mind.

This problem of identification has a long history. One still finds vestigial arrogances and stereotyped attitudes. On the one hand, there is a tendency for the psychiatrist to look upon the psychologist as a psychometrician, and, on the other, for the psychologist to regard the psychiatrist as an old-line state hospital administrator, whereas the days when these characterizations were even partially valid are long past.

Most important for the needs of society, however, is the continued emphasis on the quality of performance both at the behavioral and social levels. What we are talking about essentially is the adoption of a socially oriented attitude in the context of which psychiatry can resolve the inadequacies of its professional training, and psychology and the other behavioral and social sciences can recognize their own ultimate fundamental responsibilities toward the patient. Isn't it time to discard the overzealous *Fach* identifications of psychiatrists and psychologists which have so often been reflected in the egocentric actions of the three APA's—psychiatric, psychoanalytic, and psychological?

Some of you will undoubtedly consider this excursion into the "model" problem of psychiatry not only presumptuous, but irrelevant as well to the problem before us. I believe, however, that my four decades of participation—or if you prefer participant-observation squared—in both psychiatry and psychopathology entitles me to present the rather unique overview possible for a person "in" psychiatry but not "of" it. As to the second criticism, my point is most relevant to the conference because it has a fundamental relationship to the problem of classification. One of the great controversies in the recent history of this problem has been whether the theoretical model with which the problem of classification can

best be approached, should be a medical disease or a behavioral model.

Now that I have dealt with the two basic philosophical issues—that of balance and that of model—I wish to consider what appear to me to be six major methodological dichotomies which have to be faced in the attack on the problem of classification. These are: (1) the multiple simple as opposed to the single recondite approach; (2) the direct as opposed to the inferential approach, as reflected in both the clinical and the test approach; (3) the holistic segmental dichotomy; (4) the nomothetic as opposed to the idiographic approach; (5) the literal as opposed to the metaphoric approach, whether in the aspects analogous to the shift in biological classification from the Linnaean type of empirical description to the system of more generalized evolutionary description, or the one reflected in the structural/dynamic descriptions introduced in large part by psychoanalysis, as opposed to the older tradition of empirical surface descriptions of symptoms; and finally (6) the emphasis on method as opposed to the emphasis on content.

In approaching any problem, the investigator has the choice of developing either a complicated multifaceted program which represents a complex fusion of methods and philosophies—what I might call the single recondite approach, or he can simultaneously attack the problem through a set of separate hypotheses or methods, each of which contributes to the total picture, but is relatively independent in original execution—what I call the multiple simple approach. I believe the second approach is the more fruitful.

The complexity of the problem cannot be denied. It seems to me, therefore, that the diagnostic problem at its present stage of investigation calls for multisimple rather than a single recondite approach. Almost from the time I started my work at Worcester, I have been convinced that problems of mental disorder, for instance schizophrenia, have to be dealt with on the basis

of a group of multiple, relatively simple hypotheses. I was, therefore, most gratified to come across an important early paper on this general topic, "The Method of Multi-working Hypotheses," by Chamberlin, the famous geologist, originally written in 1890, but recently reprinted in "Science" (8). I believe Chamberlin would have accepted my extrapolation of his argument for the use of multiple simple hypotheses, to the use of multiple simple approaches which the problems of classification call for. These are the clinical psychiatric approach toward personality and behavior, the ward observer's behavior rating approach, the clinical psychologist's approach through tests and experiments, and the biochemist-physiologist's organic approach. It is only through their simultaneous use, and the combination of these various approaches, that we will eventually develop a reasonably satisfactory classification system. When the data provided by these approaches have achieved an acceptable level of dependability, there is the further possibility for the use of computer techniques to deal with their separate and combinatory complexities.

The next methodological problem is that of the direct as opposed to the inferential approach. Anyone who has worked in this field recognizes the importance of inference in the mental status evaluation of patients, no matter which approach is employed. It is true, however, that some approaches to clinical evaluation depend to a far greater extent on inference than do others. Since mental status evaluation is very rarely as objective a task as reading points on an instrument dial (and psychology warns the investigator about numerous sources of error even here), checks are called for in the area of direct observation of human behavior. Where inference plays a substantial role in the diagnostic process, the importance of controls becomes increasingly important. In addition to such controls as the use of original, agreed-upon criteria, and the use of independent judges and rating

scales for direct observation, what is additionally necessary for the more inferential methods are both explicit and implicit norms, and repeated testing and cross-checking by several competent judges of hypotheses about the latent meanings of the manifest behavior. This is true both of the dynamic hypotheses originating from clinical practice and those derived from more controlled test performances, such as projective devices.

It is important to recognize that we are dealing with an area in which the investigator or the therapist is the major instrument and we must, therefore, do everything possible to make this instrument as effective and as accurate a device as possible. It is impossible to discuss all of the problems which are involved, but in my description of the Worcester State Hospital study I have pointed out at least some of these difficulties.

Still another methodological problem is that created by the holistic segmental dichotomy. One aspect of this problem has often been characterized in medical literature as the differences in approach between the advocates of the Hippocratic approach to disease as opposed to those holding to the Platonic ideal. Henry Cohen, for instance (10), has characterized the Hippocratic view with the phrase: "*Il n'y a que des malades*," and the Platonic view with: "*Il n'y a que des maladies*." Interpreted in this way, it is difficult not to agree with the segmentalists and take the side of the *malades* in their recognition of the importance of the individual patient, as opposed to viewing disturbances as holistic entities in the Platonic sense. (One is reminded here of the old problem raised in the early days of gestalt psychology—the impractical notion of the whole world as being the only true gestalt!) Classification, however, calls for a subtle balancing of the holistic and the segmental so that the diagnosis of mental disorders recognizes the important individual and type variances which supersede fixed categories. A

more modern approach to the holistic segmental controversy should take a central position which recognizes both the total patient as a functioning unit and the patterns of individual traits which make up his totality; and that classification does not call for absolutes but rather for preponderances in both traits and degrees of their presence.

Associated with the next dichotomy, of the nomothetic as opposed to the idiographic approach, is the troublesome problem of the single case as opposed to a group of cases. Allport (3) has succinctly expressed this conflict in his question: Should scientific law be taken to refer to "any uniformity that is observed in the natural order" or should it be considered to involve only statements of "invariable association common to an entire class of subjects?" If prediction in science must, by definition, involve prediction across many individuals, then it is clear that the single case cannot be considered proper or adequate subject matter for scientific investigation. However, there have been stout representatives of the point of view that the individual may have his own laws. If we accept Kluckhohn's and Murray's (31)⁸ neat characterization of an individual's personality as being comprised of some characteristics like all other persons, some like some other persons, and some like no other person, then laws in psychological science have to take account of phenomena at these three levels: (1) universal, (2) type, and (3) individual. It is in relation only to the last level that controversy arises in our present context.

The dichotomy of the nomothetic and the idiographic (3) is only a more elaborate way of stating this fundamental problem. The nomothetic view calls for a discipline with uniform general laws, whereas the idiographic calls for a discipline concerned

⁸ This characterization was first made in an article by Kluckhohn and Mowrer (30). Rosenzweig (52) also has some points to make about norms relevant in the present context.

with particular events or particular individuals. Under certain limited circumstances, empirical prediction which does not involve intuition is possible for the individual. Psychology would, I believe, gain from the use of both of these approaches.

The characterization by Kluckhohn and Murray is relevant not only to the problem of the nomothetic as opposed to the idiographic approach, but to the classification problem as well. It is clear that the problems of classification, while not relevant to Kluckhohn-Murray's first category, the universal, are definitely related to the second category, which deals with the basic problem of typology. Although classification has only a secondary relationship to the last category in which the person is an idiosyncratic being like no other person, it is important that, in the process of working out material for the second group, we make careful studies of individuals to enable us to clarify the differences between type and idiosyncratic characteristics. It is only through accumulated intensive and careful studies of individuals that we will eventually be able to devise reliable categories of personality types.

We have already considered some aspects of the literal-metaphoric dichotomy, particularly in the analogy I used of the shift from the Linnaean empirical description to evolutionary generalization as a basis for classification in biological taxonomy. I wish here, however, to emphasize a more direct aspect, that of the conflict between the literal-superficial descriptive as opposed to the structural dynamic descriptive approach to the classification of mental disorders. I have briefly touched on this topic in the earlier discussion on dialecticism. One group indulges in a careful concern with minute discriminations and old-fashioned taxonomy in such detailed descriptions of psychiatric phenomena as seen in the studies of Langfeldt (33). The opponents of detailed description of minutiae and rigid categorization often

place excessive emphasis on dynamics. This is to be found particularly in the work of young psychoanalysts. Those who have had experience with psychiatric residents who are in analytic training are frequently distressed by some students' utter lack of interest in, or regard for, the achievement of a clear, definitive description of the patient with whom they are dealing. Instead they reveal an infatuation with psychoanalytic interpretations of the dynamics of their patients many of which appear to be spun out of illusory webs far removed from the actual personality dynamics at work in the patient himself. We thus find a contradictory situation: On the one hand, we are so engrossed with the myriad separate traits of the patient that we cannot see the forest for the trees, and, on the other, there is so little emphasis on the actual symptoms and syndromes exhibited by the patient that one has just forest—often an imaginary forest of psychoanalytic hypotheses—without any trees visible at all. What the situation calls for is a combination of both of these approaches. The generalized interpretation of the dynamics of the patient's situation is important only when it is based on accurate description of the observed phenomena. Observation and description of the phenomena must precede generalizations, the core from which dynamic interpretations can be inferred.

It is in this context that one might consider how the Karl Menninger-Mayman-Pruyser book on "The Vital Balance" (44) contributes to the solution of the problem. I would first like to pay tribute to the excellent outline of the classification systems of mental disorder which have been proposed throughout the history of psychiatry found in the appendix of the volume. It is of inestimable value to students in the field. In fact, the heroic task of these authors in the compilation of the list can only be characterized by the Hebrew term *mitzvah*, that multifaceted word which implies not only a praiseworthy moral

achievement, but also a substantial beneficial contribution in its own right. More important, the forthrightness and clear-cut stand on dynamic evaluation which permeates the book will, I trust, prove most helpful in clarifying these issues and facilitating the eventual solution of the problem of classification. Despite Menninger's nihilistic attitude toward current classification—a nihilism which I share only in part, and which I do not think should be taken literally to include all classification—the book is of immense value because of the extremely rich clinical experience on which it is based. This emphasis on clinical experience is characteristic of the Menninger Clinic's approach, whether represented by psychiatrists, psychologists, or other members of its staff.

Each of these relatively simple approaches calls for the same set of fundamental formal principles. The most basic of these is the accurate description of the patient's syndrome. This description should be made by experts who are trained in the particular approach and who have had considerable contact with the individual patient. Of the several methods of investigation, I consider, at least for the present, the psychiatric-clinical approach the most valuable. A tremendous body of important data, judgments, evaluations, and generalizations is inherent in sophisticated clinical experience. It is, of course, superfluous to mention the masters Kraepelin and Eugen Bleuler in this connection. Such clinicians closer to home as Lewis Hill⁹ and Margaret Gerard—I have deliberately limited myself to clinicians no longer alive but with whom I have had considerable contact—seem to be human psychiatric computers who arrive at mental status evaluations and diagnostic impressions based on very complex calculating processes, many of which they themselves often cannot explicitly describe. This is, I

suppose, what some clinicians mean when they claim they can "smell" a schizophrenic. Many hypotheses for experimental investigations of psychiatric and psychological questions can be deduced from the insights of these richly observant and "intuitive" diagnostic computers. While I was at Illinois, for example, we derived many invaluable suggestions for testable hypotheses about psychosomatically disturbed children from our persistent questioning of Margaret Gerard to make her attempt to explain her clinical judgments and generalizations.

I would, of course, expect the psychologists involved in classification to know their tests and experimental procedures thoroughly. But I would also expect them to be able, through general observation, to evaluate the nature of the reaction, the cooperation, and the motivation of the patient by both clearly defined diagnostic standards and through the intuition that comes from long acquaintance with mental patients.

A major contribution to the accurate description necessary for the development of a category of mental disorders is through the use of ward observers. Ward observation has, in general, been neglected. We have tended to depend either on formal tests or, more particularly, on the clinical impressions of psychiatrists who see the patients for relatively short periods of time. For more complete coverage, we have too often relied on nurses and attendants to provide us with details about the patient's behavior. Although such persons are in close contact with the patients for extended periods of time, they can only observe them sporadically and they also lack the systematic approach of trained ward observers. Ward observers should be persons who have had considerable experience in the task and who have been checked for expertness, accuracy and clarity.

The same kind of problems arise with the organic aspects. We must have personnel capable of giving the organic tests with

⁹ Cf. for instance, his book on schizophrenia (19).

great exactness while understanding the special need for careful experimental conditions with patients who, besides other difficulties, are often not cooperative or only passively so. In this context, I cannot refrain from relating an incident which occurred at Worcester many years ago during a study of basal metabolism. For a period of time—as part of our experimental caution—basal metabolic readings were taken both on the ward by nurses and in the biochemistry-pathology laboratory by a technician. We found a discrepancy between the patients' readings in the two locations. We could not quite understand why the laboratory readings were consistently higher since the instructions for the preparation of the patients and the use of the instruments were exactly the same and we were sure that the female technicians were following established procedure carefully. It seemed necessary, therefore, to instruct the internist to scrutinize carefully how the examinations were administered in both places. He found nothing in the execution of the BMR test on the ward that would account for a lower reading. When he went to the laboratory, however, he was suddenly struck by a different situation. After meticulously observing the necessary precautions with regard to the equipment and the placing of the patient in the proper position, the technician sat down opposite the patient and rested her feet on a table facing him. She was not quite as discreet about her dress as she might have been. Apparently schizophrenics do have libidos! When this "minor" factor was corrected, the readings taken in the laboratory became essentially the same as those on the ward.

But let us return to the task at hand! Our discussion of the wealth of clinical experience actually leads to the last of the dichotomies I have included: that contrasting method with content, especially in the difference between the statistical and the psychological techniques.

The best illustrations I can give of the conflict is the difference between investigators who enter the field of classification with little or no background in psychopathology, but who merely possess a special statistical technique,¹⁰ usually a form of factor analysis, or a narrow experimental technique such as operant conditioning. These would-be saviors, despite their lack of content knowledge about mental disorders, are not too infrequently off-with-their-headers, prepared not only to decapitate the poor muddled psychopathologists who have wrestled futilely with the problems of their field without being able to understand or to categorize successfully the phenomena they observe, but also to rescue the field single-handedly with the magic of their statistical or empirical methodology.

I am, of course, far from objecting to the application of new methods. In fact, the field can advance only if such new methods are continuously being tested. Method is of value, however, only when content and some understanding is already available to the user of that method. It appears most unrealistic as well as rather arrogant for "Johnny-come-latelys," without knowledge of psychopathology or awareness of the tremendous complexities involved, naively to enter this field, and expect to revolutionize it with sudden, significant solutions.

It is in connection with this topic that I wish to pay tribute to Helen Sargent, a major figure in the Menninger psychotherapy research project, who was the possessor of a wealth of clinical experience which she contributed to the general field of clinical research of which the classification problem is only a part. Her major achievement was the emphasis she persistently placed on the importance of intrapsychic events as they are reflected in the judgments of clinicians, and how these subjective data might be objectively evaluated.

¹⁰ Particularly those who in Crozier's (11) terms, use "statisticoid jugglery."

Her further development of quantitative methods, such as that based on the method of paired comparisons which she reconciled with the clinical approach, were major methodological contributions.¹¹ She suggested (54, p. 107) that criteria other than reliability and validity be used to evaluate the empirical status of clinical observation; criteria such as "reality" to be tested by agreement among observers; "relevance," to be established by the relationship of observations to other known facts; "import" thought of in terms of the hypotheses suggested by a given observation; and "utility," as gauged by hypotheses supported or confirmed. These concepts are all of great relevance to clinical research and should be considered for their ultimate significance to the classification problem.

Before closing this section I must, however, admit that knowledge can at times be handicapping, especially when long acquaintance with the field, as it sometimes does, leads to fixed and unalterable views, or when it leads to the kind of situation about which Wenkebach made his famous remark: "*Das Wissen verdrängt das Sehen*" (15). However, it is gratifying to learn that what Wenkebach was referring to in this use of the term *Wissen* was the kind of situation I have described earlier: the dangers inherent in the early acquisition of technical terms, and how such *Wissen* can act as a barrier to accurate observation.

Conclusions

In concluding, the question is whether a positive approach to the problems of classification can be culled from my extensive discussion of the difficulties inherent in the task. I have alluded to various suggestions in my discussion of the dichotomies involved. The approach which I believe will

prove the most profitable combines the following points: the use of multiple-simple approaches, care in the use of controls with either direct or inferential studies, a proper balance between the holistic and the segmental approaches, stress on the nomothetic based on extensive idiographic studies, an increased emphasis on metaphoric as opposed to literal approaches both in the use of generalizations and in the use of structural-dynamic descriptions when they are based on earlier careful use of the empiric, and a concentration on the prior acquisition of content experience as background for the employment of method.

But above all I want to stress the guiding principle I learned from the Worcester studies in this area. It can be stated in just three words—Standards, Standards, Standards.

Standards of precision are particularly important in such fields as psychiatry and psychology which are located at a precarious interface of disciplines, and in which words play such an overwhelming role in the solution of problems. In the area of classification we cannot achieve our results cheaply. We need to be everlastingly concerned about standards in observation, standards in description and communication as well as standards in syndromization and theorizing.

I believe the importance of standards in approaches to classification cannot be emphasized too strongly. You will have noted that I have, therefore, been tempted to employ the argument that the Bellman, in Lewis Carroll's sagacious¹² nonsense poem "The Hunting of the Snark," used to impress his crew with the veracity of his opinion: "What I tell you three times is true." Unfortunately, merely saying some-

¹¹ I might point out, in passing, that at NIMH we independently employed a similar technique of comparing each patient with every other patient in the group. In this study, which was undertaken independently of Helen Sargent's proposal, we obtained a correlation of .89 between a reaction time index and a measure of ego-strength or a global concept of mental health (60).

¹² "Sagacious" indeed! Saul Rosenzweig, who has engaged in a psychobiographic study of Carroll's writing, advises me that the "Snark" was "written after witnessing, with some horror, an attack of epilepsy." It appears that, in some Rhinish fashion, I had sensed that the "Hunting of the Snark" was a subliminal attempt of the poet to classify a mental disorder.

thing in the world of scientific investigation, even such a powerful word as "standards" repeated three times, does not result in the achievement of a goal as it might in Carroll's wonderland. But, perhaps, some of the "magic" we need for scientific advancement does lie in the hard work, the precision, and the avoidance of self-deception that my use of the word "standards" implies. The careful application of the points listed above may help us to catch a Snark—that is, create a workable categorization of mental disorders. If we neglect the importance of standards in either theory or method, I fear that we may instead catch a Boojum, and our efforts run the danger of "softly and suddenly vanishing away."

REFERENCES

- (1) ALEXANDER, F. (1940), "Psychology and the Interpretation of Historical Events." In *The Cultural Approach to History*, editor: Caroline F. Ware. New York: Columbia University Press, pp. 48-57.
- (2) ALLPORT, G. W. (1921), "Personality and Character." *Psychol. Bull.*, 18: 441-455.
- (3) ALLPORT, G. W. (1937), "Personality: A Psychological Interpretation." New York: Holt.
- (4) ALLPORT, G. W. (1940), "The Psychologist's Frame of Reference." *Psychol. Bull.*, 37: 1-28.
- (5) BORDIN, E. S. (1943), "Factor Analysis in Experimental Designs in Clinical and Social Psychology." *Psychol. Rev.*, 50: 415-429.
- (6) BOWMAN, K. M., and RAYMOND, ALICE F. (1931), "A Statistical Study of the Personality in Schizophrenic Patients." In *Schizophrenia (Dementia Praecox): An Investigation of the Most Recent Advances*, editors: G. H. Kirby, T. K. Davis, and H. A. Riley. Baltimore: Williams & Wilkins, pp. 48-74.
- (7) CABOT, R. C. (1912), "Diagnostic Pitfalls Identified During a Study of Three Thousand Autopsies." *J. Amer. Med. Assn.*, 59: 2295-2298.
- (8) CHAMBERLIN, T. C. (1965), "The Method of Multi-working Hypotheses." *Science*, 148: 754-759. Originally published in *Science* (old series), 15: 92 (1890).
- (9) COAN, R. W. (1964), "Facts, Factors, and Artifacts: The Quest for Psychological Meaning." *Psychol. Rev.*, 71: 123-140.
- (10) COHEN, H. (1943), "The Nature, Methods and Purpose of Diagnosis." *Lancet*, 244: 23-25.
- (11) CROZIER, W. J. (1928), "Tropisms." *J. gen. Psychol.*, 1: 213.
- (12) DEGAN, J. W. (1952), "Dimensions of Functional Psychosis." *Psychometric Monographs*, 6: 1-41.
- (13) EISSLER, K. R. (1965), "Medical Orthodoxy and the Future of Psychoanalysis." New York: International Universities Press.
- (14) EYSENCK, H. J. (1953), "The Structure of Human Personality." London: Methuen & Co., Ltd.
- (15) EYSENCK, H. J. (1961), "Classification and the Problem of Diagnosis." In *Handbook of Abnormal Psychology*, editor: H. J. Eysenck. New York: Basic Books, Inc., pp. 1-31.
- (16) FLEXNER, A. (1925), "Medical Education." New York: MacMillan, p. 253.
- (17) GUERTIN, W. H. (1961), "Empirical Syndrome Groupings of Schizophrenic Hospital Admissions." *J. clin. Psychol.*, 17: 268-275.
- (18) HAMILTON, M., SMITH, A. L. G., LAPIDUS, H. E., and CADOGAN, E. P. (1960), "A Controlled Trial of Thiopropazate Dihydrochloride (Dartalan), Chlorpromazine and Occupational Therapy in Chronic Schizophrenics," *J. ment. Sci.*, 106: 40-55.
- (19) HILL, L. B. (1955), "Psychotherapeutic Intervention in Schizophrenia." Chicago: University of Chicago Press.
- (20) HOSKINS, R. G. (1946), "The Biology of Schizophrenia." New York: Norton.
- (21) HUMPHREYS, L. G. (1962), "The Organization of Human Abilities." *Amer. Psychol.*, 17: 475-483.
- (22) HUNT, W. A., WITSON, C. L., and HUNT, EDNA B. (1953), "A Theoretical and Practical Analysis of the Diagnostic Process." In *Current Problems in Psychiatric Diagnosis*, editors: P. H. Hoch and J. Zubin. New York: Grune & Stratton, pp. 53-65.
- (23) HUXLEY, J. S. (1940), "Introductory: Towards the New Systematics." In *The New Systematics*, editor: J. S. Huxley. Oxford: Clarendon Press, pp. 1-46.
- (24) JELLINEK, E. M. (1937), "Some Uses and Abuses of Statistical Method in Psychiatry." *Biomet. Bull.*, 1: 97-108.
- (25) JELLINEK, E. M. (1939), "Some Principles of Psychiatric Classification." *Psychiatry*, 2: 161-165.
- (26) KASANIN, J., and ROSEN, Z. A. (1933), "Clinical Variables in Schizoid Personalities," *Arch. Neurol. & Psychiat.*, 30: 538-566.
- (27) KAUFMAN, M. R. (1951), "Psychoanalysis in Medicine." *Bull. Amer. Psychoanal. Assn.*, 7: 1-12.
- (28) KEMPF, E. J. (1915), "The Behavior Chart in Mental Diseases." *Amer. J. Insan.*, 71: 761-772.
- (29) KLEIN, E. B., and SPOHN, H. E. (1962), "Behavioral Dimensions of Chronic Schizophrenia." *Psychol. Rep.*, 11: 777-783.

- (30) KLUCKHOHN, C., and MOWRER, O. H. (1944), "Culture and Personality: A Conceptual Schema." *Amer. Anthropol.*, New Series, 46: 1-29.
- (31) KLUCKHOHN, C., and MURRAY, H. A. (1953), "Personality Formation: The Determinants." In *Personality in Nature, Society and Culture*, 2d ed., editors: C. Kluckhohn and H. A. Murray with the collaboration of D. M. Schneider. New York: Alfred A. Knopf, pp. 53-56.
- (32) KUBIE, L. S. (1954), "The Pros and Cons of a New Profession: A Doctorate in Medical Psychology." *Texas Rep. Biol. & Med.*, 12: 692-737.
- (33) LANGFELDT, G. (1937), "The Prognosis in Schizophrenia and the Factors Influencing the Course of the Disease." *Acta psychiat., Kbh.*, Supplement 12-13.
- (34) LORR, M. (1961), "Classification of the Behavior Disorders." *Ann. Rev. Psychol.*, 12: 195-216.
- (35) LORR, M., KLETT, C. J., and MCNAIR, D. M. (1963), "Syndromes of Psychosis." New York: Mac-Millan.
- (36) LOWREY, L. G. (1919), "An Analysis of the Accuracy of Psychopathic Hospital Diagnoses." *Bull. Mass. Com. ment. Dis.*, 3: 67-85.
- (37) MALAMUD, W., and SANDS, S. L. (1947), "A Revision of the Psychiatric Rating Scale." *Amer. J. Psychiat.*, 104: 231-237.
- (38) MARZOLF, S. S. (1945), "Symptom and Syndrome Statistically Interpreted." *Psychol. Bull.*, 42: 162-176.
- (39) MAY, J. V. (1931), "The Dementia Praecox-Schizophrenia Problem." *Amer. J. Psychiat.*, 11: 401-446.
- (40) MAYER-GROSS, W., SLATER, E., and ROTH, M. (1960), "Clinical Psychiatry." Baltimore: Williams & Wilkins.
- (41) MAYR, E. (1958), "Behavior and Systematics." In *Behavior and Evolution*, editors: Anne Roe and G. G. Simpson. New York: Yale University Press, pp. 341-362.
- (42) MEEHL, P. E. (1959), "Some Ruminations on the Validation of Clinical Procedures." *Canad. J. Psychol.*, 13: 102-128.
- (43) MELDMAN, M. J. (1964), "A Nosology of the Attentional Diseases." *Amer. J. Psychiat.*, 121: 377-379.
- (44) MENNINGER, K., MAYMAN, M., and PRUYSER, P. (1963), "The Vital Balance." New York: Viking.
- (45) MONRO, A. B. (1954), "A Rating Scale Developed for the Use in Clinical Psychiatric Investigations." *J. ment. Sci.*, 100: 657-669.
- (46) MOORE, T. V. (1933), "The Essential Psychoses and their Fundamental Syndromes." In *Studies in Psychology and Psychiatry*. Vol. 3, editor: T. V. Moore. Baltimore: Williams & Wilkins.
- (47) OVERALL, J. E., and GORHAM, D. R. (1963), "A Pattern Probability Model for the Classification of Psychiatric Patients." *Behav. Sci.*, 8: 108-116.
- (48) PLANT, J. S. (1922), "Rating Scale for Conduct." *Amer. J. Psychiat.*, 78: 547-572.
- (49) RABINOVITCH, M. S., and PHILLIPS, L. (1958), "Social Role and Patterns of Symptomatic Behaviors." *J. abnorm. soc. Psychol.*, 57: 181-186.
- (50) ROCKLAND, L. H., and POLLIN, W. (1965), "Quantification of Psychiatric Mental Status." *Arch. gen. Psychiat.*, 12: 23-28.
- (51) ROSENZWEIG, S. (1938), "The Basis for the Improvement of Personality Tests With Special Reference to the M-F Battery." *J. abnorm. soc. Psychol.*, 33: 476-488.
- (52) ROSENZWEIG, S. (1950), "Norms and the Individual in the Psychologist's Perspective." In *Feelings and Emotions*, editor: M. L. Reymert. New York: McGraw-Hill, pp. 327-335.
- (53) SANTAYANA, G. (1913), "Winds of Doctrine." New York: Charles Scribner's Sons, p. 7.
- (54) SARGENT, HELEN (1961), "Intrapsychic Change: Methodological Problems in Psychotherapy Research." *Psychiatry*, 24: 93-108.
- (55) SHAKOW, D. (1945), "The Nature of Deterioration in Schizophrenic Conditions." *Nerv. ment. Dis. Monogr. No. 70*: 1-88.
- (56) SHAKOW, D. (1949), "Psychology and Psychiatry: A Dialogue." Part I and part II. *Amer. J. Orthopsychiat.*, 19: 191-208, 381-396.
- (57) SHAKOW, D. (1960), "Modes of Adaptation with Special Reference to Mechanisms of Defense." Unpublished manuscript.
- (58) SHAKOW, D. (1962), "Segmental Set: A Theory of the Formal Psychological Deficit in Schizophrenia." *Arch. gen. Psychiat.*, 6: 1-17.
- (59) SHAKOW, D. (1963), "Psychological Deficit in Schizophrenia." *Behav. Sci.*, 8: 275-305.
- (60) SHAKOW, D., ROSENTHAL, D., LAWLOR, W. G., and ZAHN, T. P. (1960), "The Relationship of Some Aspects of Mental Set to Degree of Schizophrenic Disorganization." *J. Pers.*, 28: 26-38.
- (61) SHAKOW, D., and ROSENZWEIG, S. (1937), "Play Technique in Schizophrenia: II. An Experimental Study of Schizophrenic Constructions With Play Materials." *Amer. J. Orthopsychiat.*, 7: 36-47.
- (62) SIMPSON, G. G. (1961), "Principles of Animal Taxonomy." New York: Columbia.
- (63) SIMPSON, G. G. (1965), "Current Issues in Taxonomic Theory." *Science*, 148: 1078.
- (64) SOUTHARD, E. E., and STEARNS, A. W. (1914), "The Margin of Error in Psychopathic Hospital Diagnoses." *Boston Med. & Surg. J.*, 171: 895-900.
- (65) SPITZER, R. L. (1965), "Immediately Available Record of Mental Status Exam." *Arch. gen. Psychiat.*, 13: 76-78.
- (66) STAINBROOK, E. (1953), "Some Historical Determinants of Contemporary Diagnostic and Etiological Thinking in Psychiatry." In *Current Problems in Psychiatric Diagnosis*, editors: P. H. Hoch and J. Zubin. New York: Grune & Stratton, pp. 3-18.

(67) THURSTONE, L. L. (1937), "Current Misuse of the Factorial Methods," *Psychometrika*, 2: 73-76.

(68) TUKE, D. H. (1892), "A Dictionary of Psychological Medicine." Philadelphia: Blakiston, p. 229.

(69) WALLERSTEIN, R. S., et al (1958), "Psychotherapy Research Project of the Menninger Foundation," second report. *Bull. Menninger Clin.*, 22: 115-166.

(70) WILSON, E. B., and DEMING, JULIA (1927), "Statistical Comparison of Psychiatric Diagnosis of Some Massachusetts State Hospitals During 1925 and 1926." *Quart. Bull. Mass. Depart. ment. Dis.*, 11: 1-15.

(71) WITTENBORN, J. R. (1962), "The Dimensions of Psychosis." *J. nerv. ment. Dis.*, 134: 117-128.

(72) WITTENBORN, J. R. (1963), "Distinctions Within Psychotic Dimensions: A Principal Component Analysis." *J. nerv. ment. Dis.*, 137: 543-547.

(73) ZUBIN, J. (1934), "Tentative Scale for Clinical Diagnosis of Psychotic Patients." *Psychol. Bull.*, 31: 641.

Classification in Research on the Prediction of Response to Specific Treatments in Psychiatry

*Jonathan O. Cole, M.D.*¹

I. The Classification Problem

There are two major reasons for classifying patients into meaningful groups before embarking on clinical research aimed at determining the kinds of patients who do best on any given treatment. The first reason is one basic to the usefulness of the research. If you find that a given clearly defined type of patient does unusually well on a new drug then the results of the study are immediately useful to clinical psychiatrists generally. Second, from the investigator's point of view, the delineation of patients into a number of coherent and clear subgroups makes for clean research.

Specific hypotheses about the response of individual types of patients to the treatments under study can be developed and can be tested. One can stratify one's population into discrete groups in advance or one can restrict one's study to the specific groups of patients in which one is particularly interested.

In fact, the advantages of classifying one's patients into clean and meaningful subtypes in advance, before going into a study, seem so overwhelming one wonders why this is not always done. In point of fact, it usually isn't done. Patients, at least for drug studies and I think also for studies of psychotherapy, are usually picked in one of three ways: (a) They show a symptom which the drug or treatment is supposed to affect like anxiety or depression; (b) they fall within a large class of relatively heterogeneous patients like schizophrenics or patients with depressive illnesses or alcoholics; (c) the patient population is defined by exclusion; e.g., all patients coming to an outpatient clinic who are not alcoholics, are not schizophrenics, are not in trouble with the courts, and are not obviously mentally deficient.

All these methods tend to produce research populations for clinical trials which are relatively heterogeneous. One can argue that the results of such trials then have broad applicability. Studies of patient groups selected on such gross criteria have in fact led to a number of useful therapeutic findings such as "chlordiazepoxide is good in anxiety," "phenothiazines are good in schizophrenia" and, "imipramine is good in depression." I believe that all these statements are relatively true but they're also relatively imprecise. Everybody knows that some schizophrenics don't do well on phenothiazines, that some depressions don't do well on imipramine and that some anxious patients don't do well on chlordiazepoxide. There is much less clarity as to what kinds of patients do particularly poorly or unusually well on which treatments.

I think there are at least two factors lying behind the designing of treatment evalua-

¹ Chief, Psychopharmacology Research Branch, National Institute of Mental Health.

tion studies which lead to the collection of heterogeneous patient groups. The first of these is the imprecision of the hypothesis to be tested. One is usually looking at a drug or a psychotherapeutic modality which is proposed to be generally rather good for a broad class of people. This makes it hard to decide what specific kind of patient one should select for study and it is easier to take everyone who comes along within broad criteria and to hope that one will learn something, usually with a pious hope that an inspection of the data following the study will lead to a more precise formulation of the efficacy of this treatment for a particular class or classes of patients.

The other problem is that patients flow through psychiatric facilities in a disordered and heterogeneous way. A clinical investigator who defines his population in a very precise and selective manner may have a long wait between patients.

The hospitals involved in collaborative studies with the Psychopharmacology Research Branch have generally been relatively busy hospitals with a reasonable flow of patients. Even at these, however, getting 50 or 60 schizophrenic patients or a comparable number of depressed patients in 1 year with only modest exclusion criteria appears to be par for the course. There are not many psychiatric facilities in this country where one can be assured of such a large flow of patients that one could, for example, get 30 obsessive compulsive neuroses in 1 year for a specific study or 20 catatonic stupors. I know of a study of lithium treatment for manic-excitement which has been attempting to collect clearly manic patients from the entire state of Missouri with very little success, having obtained only about two such patients in 18 months. A study of outpatient neurotics in California had almost as hard a time getting anxious female outpatients free from psychosis or sociopathic disorders. The investigators had thought, prior to the study, that such patients were present in an abundant excess.

This phenomenon has been described as Finagle's third law. To wit: Once you try to study something it disappears! This may not be a research law at all, it may merely be a manifestation of a fact. Psychiatric patients are heterogeneous. If you get one patient who is a lovely clear example of something, you may have to wait several months to find another one who is even reasonably similar, much less exactly similar.

There is also the problem that clinicians are not, I believe, used to using diagnostic labels with precision and clarity. It is much easier for them to define a population crudely in advance than it is to establish hard criteria and then attempt to decide whether or not these criteria are present especially when decisions must be made rapidly about the inclusion or exclusion of a patient with respect to a particular clinical trial.

And to return to an earlier point, clinicians have generally not been sufficiently convinced that they had a question which required diagnostic precision of such a narrow sort to make this feature worth including in the study.

In short, diagnostic precision involving the subgrouping of patients into small heterogeneous categories has generally not been used in research aimed at evaluating the outcome of psychiatric treatments or in developing predictors for good or bad response to psychiatric treatments.

II. The Prediction Problem

There is another problem which lies at the root of our difficulties in evaluating treatment outcome and in developing predictors of improvement either in general or to specific treatments. That is the nature of the variables with which we are dealing and the way they affect patient outcome with or without treatment.

There are a few well-touted conditions in psychiatry where diagnosis is all; it leads to a specific treatment and the patient is

likely to do very well. These are paresis, and psychoses associated with vitamin deficiencies. Bromide intoxication may be another example. You make the diagnosis, you know what to do and you don't worry much about the patient's age, sex, marital status, prior personality adjustment, or anything else. Maybe you should. I would suspect that even in paresis, a patient with a poor premorbid adjustment will do poorly after his paresis has been treated with penicillin. But with this kind of condition one doesn't usually look for qualitative differences in outcome, one goes around killing spirochetes.

In neuroses, schizophrenia, depression, and the behavior disorders one may well have a different kind of situation: Here diagnosis is usually viewed as only one of several possible predictors of outcome. Does this mean we are using diagnosis too narrowly, classifying only on the basis of the present illness and presenting symptoms? Should a larger number of background variables be considered as parts of diagnosis? Are there a group of human dimensions which have an effect on outcome of all psychiatric illnesses in and of themselves or which have an effect which varies from one psychiatric diagnostic condition to another?

Do married patients always do better than single patients irrespective of the diagnosis? If they do, is marital status part of a classification system or is it an independent contributor to outcome?

Is premorbid adjustment, good or bad, always highly correlated with outcome? Do different kinds of premorbid adjustment predict different treatment outcome in different kinds of patients? Must one consider premorbid adjustment as one of the items of information leading to a given diagnosis or classification?

There is also the problem of the precipitating stress. Does this have an effect on outcome or response to any given treatment which is independent of the actual psychiatric diagnosis? Does the predictive meaning

of precipitating stress vary from diagnostic condition to diagnostic condition? How does one take account of the patient's milieu, his family, his work situation, etc., and their degree of supportiveness or rejection? Surely, these must have some effect on outcome which might well be independent, again, of the diagnosis.

The point I am trying to make is that diagnosis or classification, be it broad or refined, should be of value in predicting the outcome of treatment. Hopefully, different diagnostic subgroups respond differently to different treatments and given enough research and enough information one should be able to pick the best treatments for particular types of patients with considerable precision. If this is true, we urgently need to improve our diagnostic systems and refine our classifications and do a good deal of clinical research in which patients are placed, in advance, into specific diagnostic subgroups and their treatment responses are carefully and precisely studied. This is particularly necessary if patients really exist in a number of discrete diagnostic subgroups.

III. Do Diagnostic Groups Exist?

This goes back to what I gather statisticians call the mixture problem. If patients exist in a number of diagnostic categories, each relatively well separated from the other, they should be studied with detailed diagnostic classification prior to the evaluation of treatment outcome. On the other hand, if patients exist with a variety of kinds of psychopathological characteristics such as anxiety, paranoid ideas, depression, thinking disorder, phobias, hostility, and sociopathic tendencies, and if all of these are found to varying extents in all varieties of patients, specific diagnostic categories may not exist and we may instead have a large number of psychopathological features which are distributed in various ways, normal or skewed, among a totally heterogeneous group of patients. If this is the case, it

may be that considering an individual patient's attributes as predictors of treatment response may be a more potent way of looking for differences between treatments or looking for things which predict an excellent response to a specific treatment. Detailed classification of patients in advance may then help one little.

The premorbid personal characteristics of the patient pose a similar problem. If these have, in fact, a very potent effect on response to treatment which is independent of the diagnostic entity under consideration, or which spreads itself broadly through a rather heterogeneous group of patients such as all schizophrenics or all depressions, then one may be able to predict the patient's outcome on a specific treatment effectively by knowing a good deal about his previous adjustment and social situation without paying any careful attention to the exact diagnostic and psychopathological picture he presents on admission to your hospital or clinic. Again, discrete diagnosis may not be worth all the time and trouble involved.

Being a pragmatist, I tend to look for evidence which will show that one or another approach is more profitable or will show that a combination of several approaches will do best of all.

The past history of careful diagnosis as a way to finding out what kinds of treatments are particularly good for what kinds of patients is not awfully reassuring. This may well be because we have not been making our diagnoses very precisely and have not been using our data properly. On the other hand, it may be that the things that influence outcome are not included in our current systems for making diagnoses or they exist as rather general continua of patient descriptors whose prognostic effect is independent of the diagnostic classification in which they occur.

The crucial question I think for the long pull is, can we predict treatment response best by developing highly discrete and clean methods of classifying patients? Or are we

likely to do a more accurate job of predicting outcome if we consider a variety of attributes of individual patients as possible predictors of outcome and utilize these without attempting to organize patients into subgroups on the basis of such characteristics?

A third and haunting possibility is that response to psychiatric treatments, biological or possibly even psychotherapeutic, is determined by biological, genetic, or neurophysiological factors which are not even being tapped or measured in current diagnostic procedures since one does not know what to measure. If so, attempts to predict outcome may well continue to be approximate until we have found how to detect the biological defect and characterize it before we apply a given psychiatric treatment.

In the short run, it may not make much difference, operationally, whether clean diagnostic subgroups exist in psychiatry or whether a variety of clinically and therapeutically relevant characteristics are distributed unevenly across all psychiatric patients. It does not appear to me that we currently have either the knowledge or the statistical tools for handling a morass of variables—historical, behavioral, psychological, and biochemical, plus a bunch of treatment variables and outcome variables on a totally heterogeneous group of patients. It seems likely that many of the things clinicians deem important about patients and their responses to treatment probably do not relate to one another in simple additive or linear fashions. Most psychiatric phenomena are, in fact, distributed through patient populations in a very skewed manner. It is also likely that complex interactions occur which are ignored in some multivariate statistical approaches. In fact, one could make a reasonable bet that the interrelationships determining a patient's response to a treatment must be complex because simple relationships would have already become grossly apparent both to the clinician and to the statistician.

Given this situation, the argument for studying homogeneous patient groups, whether these are clean diagnostic subgroups or arbitrary portions of a clinical spectrum, is quite compelling, since holding some factors constant should make the influences of other variables on treatment response more readily detectable. If one could define five clear homogeneous subgroups within schizophrenia on a psychopathological basis and could look at the influence of other background variables on response to phenothiazines in each group, one might be able to tell more about the way such background characteristics operate than one could by pooling all the data on all the patients into a large matrix. A similar process should help in determining whether different types of patients respond to different treatments. In this context differential treatment response becomes a powerful empirical test of the value of the classification method employed. One way of looking at this kind of data, which I do not believe has been tried, would be a test to determine whether the categorization system used does, in fact, result in increased homogeneity of treatment response within each of the separate groups.

A number of ways of typing or classifying patients will be described later in this conference and some of them are reported to bear some relation to treatment response. Most of these approaches require replication. From the standpoint of the treatment evaluator it is devoutly to be desired that good, clean, reliable, simple methods for classifying psychiatric patients will be developed which will usefully predict response to various treatments and will be acceptable and meaningful both to research workers and to practicing clinicians.

The day of the computer is upon us and psychiatric hospitals are beginning to develop computer systems which could empirically classify each patient on admission and could select for the therapist the treatment most likely to produce the optimal

clinical result. The problem remains that we may not yet know the relevant patient characteristics to measure to permit useful classification and prediction. Two other problems also exist. The first is that we may know what to measure but the observer error in measuring it may weaken the system to the point where it cannot be practically used. The other is that the treatments whose effect is to be predicted are not unitary, clean things but are a complex set of influences composed, for example, of the drug, the dose, the duration, the patient's preconceptions, the hospital policies, the attitudes of the nursing staff, the personality and social background of the treating physician and, quite conceivably, the level of stress and frustration within the hospital staff. In the end, we may have to develop classifications of treatment conditions as well as of patients. In any event, the hope remains that the development of clean groupings of psychiatric patients will make the effects of treatment variables more readily apparent and will make psychiatry a medical discipline in which treatment can be prescribed with greater precision and greater assurance of an optimal result than is presently the case.

Discussion

LEON EPSTEIN, M.D., *Associate Medical Director, Langley Porter Neuropsychiatric Institute, University of California Medical Center, San Francisco, Calif., Chairman*

(Papers: "Epidemiology and Medical Care Statistics," Ernest M. Gruenberg, M.D., Dr. P.H., Professor of Psychiatry, Columbia University, New York, N.Y. "Classification of Mental Disorders for Epidemiologic and Medical Care Purposes—Current Status, Problems, and Needs," Morton Kramer, Sc.D., Chief, Office of Biometry, National Institute of Mental Health. "The Role of Classification in the Development of the Science of Psychopathology with Particular Reference to Research," David Shakow,

Ph.D., Chief, Laboratory of Psychology, National Institute of Mental Health. "Classification in Research on the Prediction of Response to Specific Treatments in Psychiatry," Jonathan O. Cole, M.D., Chief, Psychopharmacology Research Branch, National Institute of Mental Health.)

OPEN DISCUSSION

(Dr. Epstein opened the meeting to discussion and served as moderator.)

Dr. BAHN. I want to thank Dr. Gruenberg for clarifying the difference between multiple diagnoses, or multiple disorders, and the use of multiple dimensions to describe the condition a mental patient might have.

It is the latter approach the Committee on Psychosocial Diagnoses and Classification of the American Orthopsychiatric Association is interested in—the level of functioning and role performance. The committee believes that much greater agreement can be achieved about a multiple dimensional typology than a unidimensional typology with its constraints. A multidimensional scheme is more useful at the present time for clinical and research purposes. I don't think you can hope to get agreement on a unidimensional typology, as evidenced by the problems we are having with the APA classification. Problems in the use of the APA classification are particularly exemplified in data on children, where 40 percent of the cases that are terminated from clinics are either classified as transient situational personality disorders, or as undiagnosed; I don't know what you can really do with that kind of information.

I would agree with Dr. Kramer that a nomenclature and a statistical classification are different. But I wonder if some classifications could be, if properly structured, both a nomenclature and a classification. I refer, for example, to the recently developed "Standardized Nomenclature on Pathology" which serves both as a nomenclature and as a classification system. And I think there is some evidence that we do

not want to condense our terms too rapidly before we know what items are important. Actually, other than brain syndromes, the terms in the "Standard Nomenclature" that pertain to the mental disorders are also the terms and categories used in the "International Statistical Classification."

Dr. GRUENBERG. Dr. Bahn emphasized a point implicit in what Dr. Kramer said. It is dangerous to think that you know something is wrong with your classification because 40 percent of your cases fall in a single category. I don't see how the distribution of your findings tells you whether your classification is good or not. And I don't know of any serious problem in using the APA classification. You get your reports. I would like to ask: Who has problems with this? Psychiatrists? I have no problem with it.

Dr. BRILL. I think we have two topics that we are discussing together. In the first place, there is the problem of classifying those things which we know. The second problem is what we do about what we don't know. I am afraid a great deal of our problem in child psychiatry falls into the category of what we don't know. It is very hard to classify one's ignorance in a satisfactory fashion. It may be that a lot of our discomfort stems from the fact we have some spadework to do in learning more about child psychiatry.

Dr. KRAMER. I would agree with the point that Dr. Gruenberg made about the size of the category. The fact that you get a lot of cases falling into it doesn't necessarily mean the classification is not a satisfactory one. It depends on what one wants to find out and what kind of classification one needs to establish in order to answer these questions.

There was also the point made about whether you could use a nomenclature as a classification. The answer is a qualified "yes." A nomenclature has a rubric for

every conceivable disease, and from the point of view of statistical analysis of categories with a small number of cases you run into impossible situations. It might be important to know you have one case of a rare disease or of an infectious disease that has been brought under control. From this point of view it is important to know that this rare or unusual disease is occurring. If we suddenly saw a case of smallpox in this room, it would be important to know about it.

The major point, in relation to a nomenclature versus a classification, is that the classification provides a way of systematically condensing the categories in the nomenclature into broader groupings of related conditions for various types of statistical analyses of the data.

Dr. BAHN. I would like to comment further about the point that Dr. Gruenberg made.

I would agree that if a patient is classified as a transient situational personality disorder we know that this patient, if properly diagnosed, is not schizophrenic. However, that is about all we do know, because the information itself is so vague. It serves a limited purpose, from the standpoint of clinical case management, as well as for research purposes; and, therefore, its utility is limited.

If having a transient situational personality disorder connoted a lot of other characteristics with a fair degree of certainty, I think it would be useful. You know, perhaps, that the child does not have schizophrenia or a brain syndrome, but there isn't much more that you do know. I would like to have some of the child psychiatrists comment on this point.

Dr. JENKINS. As a child psychiatrist, I volunteer on this point. There is a subdivision of the adjustment reaction of childhood which does not help Dr. Bahn because it doesn't have separate numbers in the statistical scheme. But you recall there is conduct disturbance separated

from neurotic traits, and separated from habit disturbance. And the conduct disturbance versus the neurotic traits does at least give us a broad differentiation between what in the adult we call the psychoneurotic and what in the adult we call at least certain of the personality disorders, the sociopathic group.

Now I think there is remedy in this area coming as a result of the work of Dr. Brill and Dr. Pasamanick and Dr. Kramer and others, because the eighth edition of the ICD retains what was in the seventh edition—I believe it was the seventh edition—a category of behavior disorders of childhood. I am proposing that we subdivide that into five groups, because some fairly repeated and extensive work indicates there are five symptomatic clusters easily distinguishable among children. These clusters correspond, for the most part, with the major diagnostic divisions in the adult area. At this point I think child psychiatry is moving up to the level that adult psychiatry has occupied. I am suggesting the overanxious type of behavior disorder for the group resembling the neurotics, the withdrawing type for the group resembling the schizoid and the schizophrenics, the undomesticated type for the group resembling the antisocial reaction, the hyperkinetic type for the group resembling the brain syndromes. And then there is another type corresponding to a personality diagnosis omitted from the eighth edition of the ICD but retained as a descriptive term, the dissocial type.

This brings us to the point which Dr. Bahn, I believe, is necessarily concerned with: We have problems that are not problems of psychopathology, but rather are social problems. I think we have to recognize a difference between individual psychopathology and pathology of social relations and social organizations. This involves a recognition of a predatory criminal culture, for one thing. This was well exemplified and well institutionalized in the "Thugs of India," and such a

predatory criminal culture persists in some sort of fashion with us today. We need to recognize that an individual can be psychiatrically normal and very unhappy in a marital relation, even though married to somebody who is also psychiatrically normal. Somehow we must get into this area, too, and yet not confuse it with individual psychopathology.

Dr. EPSTEIN. Dr. Kramer, would you care to comment on this?

Dr. KRAMER. One point with respect to this dissocial reaction: It is in the ICD, but it is not in section 5. When you try to establish an international classification you invariably run into situations where one must make compromises. Some of our European colleagues refused to accept this condition as a medical diagnosis in a medical classification of diseases. Therefore, it is to be included in that section of the "International" which provides for non-medical conditions recorded in the medical record. Since there is a place for it in the "International Classification," it can be brought back into tabulation lists for psychiatric disorders used in the United States.

Dr. JENKINS. Did they put a normal pregnancy in the same place?

Dr. KRAMER. Yes. They have a whole series of conditions.

Dr. PASAMANICK. I would like to clarify it just a bit further; there is—or will be—a classification as one of the categories in the YY (no disease) category. I don't recall the precise term. It is stated somewhat like this: Social maladjustment without psychiatric disorder.

For a long time, by the way, normal pregnancy has been in there as one of the conditions that is classifiable but is not a disease.

The difficulty that arose with our Eastern European colleagues was in the misuse—and in this case I think most of the people present agreed that it was misuse—of the

antisocial category to categorize all delinquents and all criminals, or many criminals and many delinquents.

The antisocial personality disorder category has not been eliminated. It was hoped that our psychiatric colleagues in the West would not misuse that category if the other category were available to them.

Dr. EPSTEIN. Dr. Gordon.

Dr. GORDON. Mr. Chairman, while representing the American Medical Association at this important and most interesting session, may I mention the views and experiences of the AMA in the development and furtherance of "Standard Nomenclature of Diseases and Operations" (SNDO) during the past 15 years. You may realize that the AMA has served as custodian of SNDO since 1937 with editorial responsibilities for the development of three editions. This book has enjoyed wide acceptance, but in 1960 questions were posed about the complexity of the system, notably failures of the text to reflect the advances of modern medicine.

Detailed studies of SNDO in 1961 revealed that the basic problems were related to the plethora of names and/or terms for diseases, especially outmoded eponyms, synonyms, and generic designations as encountered in surveys of standard textbooks of medicine and the current literature. Accordingly, the difficulties seem to be rooted in the language of medicine itself. The situation was emphasized during the development of "Current Medical Terminology." For example, 18,000 different names for diseases were noted in recognized medical publications in which more than 10,000 terms were in direct conflict, thus creating confusion and difficulties in filing, evaluation, and communication.

It should be mentioned, perhaps, that the new edition of CMT will contain 3,000 preferred terms; for acceptance, each term is supported by a definition containing significant etiologies, manifestations, tests, and pathologic findings. Other terms fail-

ing to meet the requirements of accepted or preferred terms were set aside for further consideration. The new edition provides an alphabetical listing of supplementary terms, as certain eponyms and synonyms, a key-word-in-context index to illustrate the specificity of descriptors for the description of manifestations, a separate glossary of descriptors, and a numerical index. The book is computer oriented to facilitate timely revision, statistical analysis of data, and special studies in depth as for programmed teaching and diagnosis.

Dr. GRINKER. A quotation was made this morning from Dewey and Bentley; namely that "Naming is knowing." One could just as well have twisted this around and said, "Knowing is naming." The question is not the name; the question is, "How do you know?"

I think that we can use many words in our classifications. The fact is we can amplify according to many traits, as has been brought out in the discussion, which will create more names, but hardly more knowledge.

We can, of course, start out with the assumption that we have a nosological classification, and we can say that we would like to study those conditions that psychiatrists usually call X, Y, or Z—depression, schizophrenia, borderline, and so on. And for that purpose we have the general prototype of research, which has not been bettered: that of Dr. Shakow, although in many respects it has been amplified by others.

But once we have made a decision that certain groups of cases are, with a high degree of reliability, designated as a certain category, then we may analyze these categories in terms of their various traits. These traits may be determined a priori from what is in the literature, from what is communicated at conferences, from what one obtains from one's own personal experiences; or the traits may be developed

in the process of study of the various categories.

These traits may then be determined in terms of their frequency, and their extent quantitatively, by observations of behavior, by interpretative inferences, by verbal communications, and by a wide variety of biological tests. And each one of these areas may be used separately or together. And this is, in general, I think, what Dr. Gerard was doing in his schizophrenic research.

And from this one may crystallize out types.

So, for example, in determining what is schizophrenia, working with Dr. Sam Beck, we decided on a definition of schizophrenia that we could agree on. We called it our definition. When he isolated the so-called six schizophrenias he did not give them names; he said, "These are the six schizophrenias, the types we have been able to find." We could then test out the validity of these categories with a good degree of reliability by the process of correlating them with other systems.

We may use the system of prognosis, the system of therapy, the system of biological factors, or any other system for validation. But when we come out of this kind of research, we are properly able to amplify the classifications.

I think this is the kind of research that Dr. Shakow started; and again I say, the prototype of all good research is along these lines.

Dr. SHAKOW. I might say I agree heartily—not with the comments about my research, but with the method.

We certainly were not trying to get a sample of schizophrenics—that is, the schizophrenics—in Worcester State Hospital, or those who came to us from the Boston Psychopathic Hospital, or elsewhere; we were interested primarily in studying schizophrenia. That meant we had a lot of criteria. We said: "Don't take anybody over 50 years old; don't take any

women. (This was not because we were prejudiced against women, and we recognized that actually they would perhaps be much more "interesting" cases to study. Since many of our studies were endocrinological in character, however, it seemed better not to complicate our problem at that time with the study of women. We did study women afterwards for other special reasons.) But we followed certain selective criteria—those I have mentioned. Thus, we didn't mix in alcoholics in our group. Further, in the process of diagnosis, having established clear criteria, we were careful about keeping our subtypes clear.

First, for instance, we were sure—as sure as our study group could be—that we had cases we could call schizophrenia. When, for instance, we called a patient catatonic, he had to meet certain kinds of criteria on which we were all agreed. If he had a mixture of other types with the catatonia we called him mixed.

I think it is by getting down to a relatively few, clear-cut psychiatric diagnostic criteria that you can begin to make comparisons with biological studies, with psychological studies, with social or cultural studies, or any other approaches.

Obviously, if somebody said, "Are you studying schizophrenics as they exist?", we would have had to say, "No," since so many schizophrenics have these mixed qualities.

But we weren't trying to answer that kind of a question. We were trying to parcel out what was essentially schizophrenic, or what was characteristic of a particular kind of schizophrenic.

One of the problems that has existed, especially in recent years, has been that it has been most difficult to get "clean" cases. Luckily we were working, at that time, before the insulin days, before the metrazol days, and before the shock therapy days. The kind of therapy which our patients had was a kind of custodial social therapy—not a particularly distorting kind of therapy! And life was relatively easy.

I know the problems we have at the present time in trying to get good material to work with. And I am sure that it would be very difficult to repeat studies such as we made in the late 1920's and the 1930's at the present time.

But I think that, somehow or other, there has to be a series of studies of relatively small but intensely studied groups. From such work we will eventually get some kind of system which would perhaps prove satisfactory.

Dr. KLERMAN. I read Dr. Shakow's paper with much admiration. Yet I am struck by the discrepancy between his excellent description of the Worcester project and some of his comments about the medical orientation of psychiatric research. The work at the Worcester Foundation stands as a model for interdisciplinary work such as has been carried on by Dr. Gerard, and also to a certain extent by Dr. Kety, as well as others. Unfortunately the recent work has not produced any more information about the basic etiology of schizophrenia, although their methodologies are in some ways more refined.

Having reviewed these things, Dr. Shakow, you then say that you feel the medical orientation of psychiatry should be questioned. I question what would follow in terms of research design, or research strategy, if we took your admonition to heart.

It seems to me that your own work during that period—and some of the recent work following that model—is very much within my understanding of the grand tradition of medical research; that is, the investigator starts with the clinical phenomena, defines its limits, attempts to find correlates (in the hope that finding correlates of clinical phenomena with independent phenomena will elucidate the etiologic processes). And yet, having done this kind of work, you now say that you question its philosophic basis and its his-

torical tradition. I would appreciate clarification.

Dr. SHAKOW. Of course, I think what you said last was really a description of any kind of scientific work, medical or non-medical. Any kind of research work calls for principles of the kind you describe.

I was questioning whether psychiatry was not somehow different from other medical disciplines. It happens that historically psychiatry grew out of medicine, because the early superintendents of hospitals, for one thing, were medical people. This was the only tradition in which it could grow up.

As I have examined and watched psychiatry over four decades, I have been impressed that, generally, psychiatry has not been concerned with what would ordinarily be called medical problems. It has been concerned primarily with problems of behavior.

I don't have any prejudices with regard to what psychology is—that is, I think that different persons may be psychologists, depending upon their training and the kinds of things they do. I don't care what background they have, as long as they get the proper kind of training. So a psychiatrist can be a psychologist, and a sociologist can be a psychologist—if he wants to study behavior, and if he goes through the necessary training.

Since I don't have that kind of prejudice, I ask: Why is it that psychiatrists use a model which may not be the most appropriate model for them—that is, the disease model—when they might be able to work out, on a basis of personality differentiation, a model which is much more useful for them than the disease model?

We have been hearing today about the use of the medical disease model, but not too much discussion of a behavioral model of some kind. I deliberately and provocatively put the question in this way to raise this issue: Is psychiatry actually using the proper model?

I think people who are going into social psychiatry recognize the need for a behavioral model. I think some of the people who are concerned with biochemical aspects, perhaps, and some of the physiological aspects, have more rationale for, as I said, using both kinds of models. Psychiatry, it seems to me, is in a particularly difficult spot here.

Careful thought should be given as to the nature of the model which is most appropriate for classification systems in mental disorder. I don't think it should necessarily follow the present model, which is a medical one.

Dr. GERARD. I find the tenor of the discussion is now developing very much to my taste. I am, therefore, unable to resist taking the floor, although I was going to wait until the session when my paper is discussed.

I want to disagree in detail, although I think not in principle, with the last thing Dr. Shakow said. I wonder if this is a helpful way to look at things.

One has to classify, otherwise one has an amorphous universe with which one cannot deal intellectually or operationally. One does classify, in general, for one of two reasons, which were enunciated in the opening paper: either to do something about the different groups so named, or—and this hasn't been quite so clearly brought out—one classifies on the basis of, and to promote, understanding.

We have talked as if the problem in classification were primarily how you grouped various kinds of units. I submit that the main problem in classification—and certainly in the understanding aspect of it—is finding the right kind of units to classify, the thing I have called entitation of the universe.

Any amount of study of the wrong entities is rather useless, and often very misleading.

One is faced, practically, with classifying certain phenomena seen in certain human beings. And one has to do something with

them, and one makes names and one makes systems. As has been made abundantly clear, these are not very enduring because, up to the present—certainly in the mental area—they have not been based on very profound understanding.

The hope is that by the kind of work which is going on, and has gone on, one will be able to, so to speak, dissect out entities which are meaningful.

Now “meaningful” to me—and perhaps this is merely revealing a prejudice—connotes that one has found a real entity, not because you have called it this or that, but because it is a valid one.

The extreme problem about disease, more in the mental than in the ordinary physical area—though of course they are the same, really—is that these are functional units rather than material units. It is much easier to identify material units and to group material units than it is to identify meaningful functional units and then group them.

There is not much trouble in telling an individual animal from the world around it, and recognizing it as such. And there is not much difficulty in seeing certain relationships between the individual domestic cat and cats that look like lynxes or tigers or lions. These are relatively simple problems, first, of identification or entitation, and then of classification.

But we are faced with the identification and then classification of functional units. Disease does not exist as such. Disease is a functional condition in a material unit. The essential problem there is to regroup the approach to it from looking at the material units as your entities, to looking at the material units as the bearers of the entities that you are interested in studying, or that we are interested in studying.

This is enormously difficult. Nobody need bend his head in shame in a field where we are limited primarily to behavioral manifestations. I don't think that Dr. Shakow's discrimination is really an important one. Whether you talk in the mode of disease or whether you talk in the mode of behavior units, you are in each case talking about a functional situation, and it is meaningful to give it some kind of a name if, indeed, it turns out ultimately to have a wholly consistent, reasonably coherent, and extensive group of phenomena that can be so characterized.

In the past, I have often used the example of the nerve fibers: nerve fibers are very nice, clean, material entities that were discovered very long before the nerve impulse. The nerve impulse is a functional unit, however, which has just as precise a meaning; it is just as easily identifiable, just as fully describable as the nerve fiber. You can't experiment with the nerve impulse; you experiment with the nerve fiber. But the nerve impulse has certain electrical, chemical, temporal, functional attributes that go together, and there is not the least doubt in anybody's mind that this is a meaningful entity, as a reflex is, a conditioned reflex is, and so on up.

I think our problem is that we still have not really found—or at least if we have we are not sure whether we have or have not found—the really meaningful entities of behavior or of disturbance, the functional entities that underlie real, meaningful mental illnesses. When we have them, then I have not the slightest doubt myself—again, perhaps a prejudice—that these will correlate with etiological and prognostic and therapeutic and other aspects which will make them utterly meaningful and very satisfactory.

C. The Role and Problems of Classification in Related Disciplines and Their Relevance to Problems in Psychiatry and Psychopathology

The Role of Classification in Personality Theory

*George A. Kelly, Ph. D.*¹

A theory is a device to enable man to deal comprehensively with what might otherwise be an overwhelming variety of events. Some of the events may not even yet have occurred. Moreover, there are myriads of them—some fascinating, some propitious, some perhaps catastrophic—that may indeed never occur unless a theory prompts someone to undertake them.

Strictly speaking, man invents his theory; he never discovers it. What he discovers, often to his dismay, is how well the theory works. But even that, as I suspect every venturesome theorist is painfully aware, often turns out to be a matter of taste and momentary convenience. But whether or not it works to suit its users, the theoretician must hold himself responsible for what he has contrived. Much as he might like to shift the responsibility, he cannot ever claim that his theory was dictated by the facts it seeks to explain. The facts tell him only when he is wrong—and not always that. And besides, there simply are too many ways of explaining the same facts—including a lot of ways that haven't turned up yet—for any of us to claim privileged communication with either God or Nature.

The process by which a theory is devised is known as abstraction, and the complementary process by which it is extended to events is known as generalization. Both may properly be regarded as psychological processes, or even as aspects of the same

psychological process. That is to say, to understand them one may, instead of relying upon the principles of logic, employ some notion of man's style of life. Thus meaning is not extracted from nature, but projected by man upon it.

Neither abstraction nor generalization has ever been computerized, nor can either be realized by any unimaginative obedience to the canons of rationality, or by performing the symbolic transformations of mathematics, useful as these procedures may otherwise be. What can be computerized, for example, is the elimination of redundancy in a construction matrix. The resultant shrinkage in the matrix is sometimes mistaken for abstraction, for it appears to result in the expression of a great deal in relatively few terms. But the contribution the computer makes is to the economy of the language employed, not to conceptualization; albeit one must grant that linguistic parsimony may serve to clear away the clutter that stands in the way of fresh thinking. But housecleaning is not abstraction, and economizing does not constitute theoretical thinking, Ockham's razor is a surgical instrument, not a creative tool.

1. A Construction Matrix

By a construction matrix I mean a postulated grid in which events and abstractions are so interlaced that whatever appears to occur independently of one's intention is given meaning in depth by being plotted against whatever coordinate reference axes

¹ Professor of Psychology, Brandeis University, Waltham, Mass.

he has intentionally erected. And in this psychological hyperspace the humanly contrived axes of reference, in turn, acquire whatever objective significance they have through extension—or through “operationalizing,” if one prefers a term that has more current usage.

This is to say that human constructions derive their objectivity wholly from the way they cast events into varying arrays—or simply from the lines of perspective they provide. Actually it is in terms of such arrays that consensual judgment becomes psychologically possible. Consensus itself, while often cited as the criterion of objectivity, does not properly define the psychological grounds on which objectivity rests. Only sociological grounds are implied.

But now, since we are talking about human experience, including our own particular experience as scientists, it may be more precise, instead of saying that the matrix is a schema in which events and abstractions are interlaced, to say it is man's observations and his constructs that are woven into the fabric of experience—the one ascribing meaning to the other and the other lending palpability to the one. And in this more phenomenological sense the grid might better be characterized as a “repertory grid,” since it expresses one's own finite system of cross-references between the personal observations he has made and the personal constructs he has erected. I suppose it is apparent that all of us must have quite limited repertoires, for the events we encounter are experienced only in such depth as our constructions will plumb, and our constructs have only that scope which is provided by the ranges of events to which we undertake to apply them.

2. A Theory as a Psychological Venture

In these few sentences I have tried to sketch the principles that are involved when one seeks to understand a theory as a psychological venture, rather than as a logical

entity. Perhaps I have tried to say too much in altogether too little space. Obviously the farflung implications of this psychological approach to human theorizing have not yet been mentioned. Nor do I have any intention of trying to spell them all out, except as they bear upon the problem addressed by this conference—“The role and methodology of classification in psychiatry and psychopathology.”

What I have said is that a theory is a human contrivance, not something for which a man can escape personal responsibility by claiming he discovered it somewhere, that the abstraction process by which he fabricates his theory is a reflection of his own strategy for coping with events, and not a means of distilling truth out of the events themselves, that the condensation of a construction matrix—or information net, if you prefer a more limited and mechanistic notion—through the elimination of superfluous data and redundant terms is not to be confounded with abstraction, that man comes to grips with his existence by plotting his observations against his constructs and by testing his constructs against his observations, that meaning in the human role each man plays is limited by the variety of constructs he is able to devise, and that our constructions have no more substance to them than the extent of the events to which they are applied.

3. The Scientist as a Man, or Man as a Scientist

There is something interesting that happens when one examines theorizing as a psychological process. He finds himself describing his own psychological efforts in the same terms he uses to account for those persons about whom he is psychologizing. This can be disconcerting, and, depending on the amount of cynicism built into his favorite personality theory, can lead to his having some very grave doubts about the validity of his efforts.

Most psychologists get around this by using different theoretical frameworks to explain the behavior of the scientist and the behavior of the ordinary man. For example, the scientist, they will say, systematically formulates his outlook, generates hypotheses from it, boldly ventures to make predictions, undertakes audacious experiments, courageously pits speculations against outcomes, candidly observes what happens, and completes the experiential cycle by humbly revising his outlook from the fresh vantage point to which his earlier naive commitments have led him. But ordinary men are held accountable in quite different terms. They are believed to be conditioned by the events that impinge upon them, sucked into the vacuums created by their needs, propelled by drives that invade their otherwise placid lives, caught up in the wake of their cultures, locked into the orbits of their biographies—lacking the imagination to attempt anything they have not rehearsed—or kept alive and kicking only by yielding to the persistent nagging of external stimuli.

But whatever we say about the behavior of the scientist is itself a personality theory, even though we may seek to limit its application to licensed scientists, or to other classes of “white folks.” On the other hand, whatever we choose to say about human behavior in general is a personality theory that must deal not only with man’s indiscretions, but hold itself accountable also for the structure of the scientific edifice that man, and man only, is known to have produced.

4. Categories as Obstacles to Theoretical Development

What happens when we undertake either to view the scientist as a man or man as a scientist? We get into categorical difficulties. And it is to just such difficulties imposed by categorization on theoretical thinking that a conference on classification, such as this one, should address itself.

If we view the scientist as a man we must regard his science with the same incredulity we apply to human behavior in general. This is to say that most of us would then have to claim that one’s science is a symptom of his psychodynamics, or perhaps a product of his operant conditioning, though modulated, to be sure, by the psychodynamics—or the conditioning—of those with whom he is dynamically identified, or to whom he is conditioned. Thus, if we were to persist in accounting for scientists in terms of current psychological theories about men I think we would soon be led to the conclusion that science is more a human predicament than a vital undertaking. On the other hand, if we apply the scientist paradigm to man, we someday are going to catch ourselves saying, in the midst of a heated family discussion, that our child’s temper tantrum is best understood as a form of scientific inquiry.

In the one case—in the case of understanding science psychologically—we are limited by current personality theories that have never had to bridge the categorical boundaries customarily set up between natural logic and human behavior. Thus, by holding to the stubborn belief that science is human only when it is in error, we perpetuate some preposterous myths about the pure scientific enterprise. In the other case we are blocked by the ancient barriers that in Western culture seal off the classical categories of cognition, affect, and action. So we argue that since a temper tantrum has been categorized as an eruption of affect, one need accept no obligation to explain it in terms consistent with cognition. In each case, therefore, instead of using our constructs as reference axes in terms of which events can be seen in multi-dimensional depth, we continue to use them to construct pigeonholes in which events are tucked away and rendered inaccessible to other dimensions of appraisal.

But why not go all out and suggest that science, as we know it today, may be a reasonably apt—perhaps a tragically apt—

statement of this century's human predicament? Or, if we may be permitted to take science itself as the model of a personality theory, why not then regard the child's temper tantrum as a frantic experimental effort to articulate some urgent question about human relationships for which no one so far has been willing to give him a candid answer? Why not? Why, in the first place, we say, because science obviously is a logical statement about nature itself, and therefore could not possibly be a 20th century illusion created by man! Why not? Why, in the second place, we say, "a temper tantrum is an emotional outburst," and therefore clearly could not be a cognitive inquiry! Thus we see examples of the categorical barriers that block the free use of theory in the extension of human inquiry.

5. A Theory as an Aid to Creative Thinking

One of the unique functions of a theory is to enable man to reach beyond what he already knows. Having devised a system of constructs, and given it anchorage in his observations, one may state the parameters of some event he has never seen, but which might reasonably be expected to occur under certain conditions. Thus a good theory is fertile soil for producing something new, whether a new way of coping with mental disorder or a fresh approach to the education of the child. For my own part I would place fertility at the top in any list of criteria for a good theory, particularly for a good personality theory. This, however, is not to overlook the importance of coherence, relevance, comprehensiveness, parsimony, verifiability, rigor, and predictive efficiency in theory construction.

From this point of view, categorization, helpful as it may be for certain practical purposes, proves itself to be almost completely sterile in suggesting something new to be looked for. It rarely leads to experimentation, and, while at times it may challenge one's legalistic ingenuity in reaching a diagnosis, it fails to set one's imagination

on fire. It does, of course, tend to narrow down a scientist's field of inquiry, and on some occasions that may be just what he needs in order to develop a new construct more precisely within some limited range of convenience. Yet hardening of the categories, a common affliction among scientists, usually marks the end of the creative phase of a distinguished career.

6. Efforts To Escape Categorical Restraints

From time to time in the history of human thought men have been acutely aware of the barriers that categorization has raised against their efforts to do creative thinking. It is no accident that dogmatism, the mast that surrounds all bastions of classical ignorance, is characteristically categorical in its logical form.

Among the obstacles produced by categorical thinking probably none has been more frequently assailed than that imposed by the dualism of mind versus body—or the spiritual versus the material. Typically the efforts to transcend this duality have taken the form of placing everything into one category or the other, as, for example, in idealism and materialism, or by envisioning some sort of interaction between mental things and physical things, as when we speak of psychosomatic disorders. But rarely in science have we given thoughtful consideration to the implications of categorization itself, and to the occasions when it would be better to structure our scientific efforts in other ways.

But now this conference has been called, and I am sure there is more afoot here than merely another updating of the approved system of psychiatric nomenclature. The notion of mental illness has itself been seriously challenged for the first time since the days of the Nancy School. Even the continued use of the so-called medical model in psychiatry is being thoughtfully questioned. Social psychiatry has brought fresh attention to the nature of psychological processes and has given relatively little play to patho-

logical categories. Almost every new contribution to personality theory or to methods of treatment—from psychoanalysis to tranquilizers—has downgraded the importance of categorization of disorders and substituted some abstracted notions about how man copes with his circumstances. The new emphasis is upon what man does, rather than upon what happens to him—upon what he undertakes, rather than the state he is in.

7. Constructs as Reference Axes Rather Than as Denominations of Events

Up to this point I have been comparing and contrasting the aims of theorizing with the effects of categorization. I have pointed out that they usually serve different ends and that they too often get in the way of each other. But categorization, with its preemptive claim to the events it structures, is only one of the forms of classification, and since this conference has to do with the role of classification in general rather than categorization in particular, we may ask ourselves if there is not some other way to engage in classification without obstructing the development of better personality theories.

The key to our problem lies, I believe, in distinguishing clearly between the events we wish to control and the constructs we devise for understanding them. In terms of the repertory grid I have invited you to envision, this means differentiating sharply between what is warp and what is woof. This is to say that any clinical observation we make has meaning only in terms of the constructs with reference to which we choose to plot it. And our constructs, in turn, are not homologous with the events they enable us to understand, even though our language is grammatically constructed to make them appear that way.

We observe a patient's behavior. But it makes no sense until we plot it with reference to some—such notion as, say, anxiety. This does not mean that what we have

observed is of itself an anxiety. It means, instead, that if we construe it in terms of anxiety it begins to make sense. Similarly anxiety is not to be regarded as a particular collection of behaviors, but rather as a notion we have invented in terms of which the events observed in a patient's life appear to fall into some orderly array. Anxiety, then, is not a category of patients, nor is it even a category of symptoms, but it is a contrived reference axis against which any behavioral observation may be plotted, even including observations that may stand out much more clearly in the light of other constructs.

8. Propositional Rather Than Constellatory Relationships Are Required for Theoretical Advancement

In this kind of epistemology constructs are propositionally related to each other rather than bearing constellatory obligations to each other, or preemptively excluding each other. Let me explain what I mean. I observe a bit of behavior. It makes sense, it seems to me, to regard it as an expression of anxiety. That is to say I have made a proposition about its position in relation to my notion of anxiety. But does this mean that I must also regard it as neurotic? Not necessarily, for I may, if I wish, devise a construct of neuroticism which does not embrace all forms of anxiety. The two constructs thus need not be constellatorily related in my chain of inference. And certainly I do not need to insist that my diagnosis of anxiety precludes all other considerations. This would be preemptive construction, and it represents the worst of what can happen in a system of diagnostic categories—the worst, that is, from the point of view of a theoretician who is always hoping to see things in a new and better light.

Theory building relies heavily upon the propositional use of constructs. Thus the theoretician, no less than the dedicated psychotherapist, keeps himself alert to

potentialities not yet realized in a patient's life. But the more practical and decision oriented a clinician becomes the more he hastens to group his constructs into constellations, making inferences from his placement of an event on one reference axis to its probable placement on another. He might, for example, say that because his patient's behavior looms up so clearly in the light of his anxiety axis he had better go ahead and regard it as neurotic too. And in hospital management, where so frequently the most urgent decision made during the first week is to which of the overcrowded wards the patient should be sent, the clinical director is only too happy to have the fellow tentatively, but preemptively, consigned to a diagnostic category before the intake staff takes off for the day.

So far I have described the theoretician's constructs as reference axes for appraising objects, rather than as names for objects or groups of objects. And I have pointed to the theoretical advantage of using them propositionally, so that the full range of implications of each construct dimension can be explored separately without being encumbered by the implications of other constructs previously assumed to be related to it. For example, we can examine empirically all the implications arising out of the construct of anxiety, without necessarily having to concern ourselves with neuroticism.

9. The Noncategorical Use of Classification: Cluster Referents

Now it is possible to employ classification in science without resorting to categorization. Biologists, who are oldtimers in the categorization game, are just beginning to catch on to this notion, but librarians and efficient secretaries have known about it for a long time. Used in this way classification becomes a system of cross-referencing or indexing. Of course the books are physically located on some shelf

and the papers have been put in some particular folder, and that much of it is categorical. But you don't keep track of the books and papers by running from shelf to shelf and from file to file. In this sense the classes not only overlap, but cut across each other, as the terms of classification are rotated into alignment with other axes of reference.

When we use classification in this non-categorical way the classes are retained as abstractions and are not walled off concretistically to encapsulate events. Though they have the logical form of classes, they are actually used as nonparametric reference axes. A particular set of behavior observations, for example, can then be described all at once as anxious, fragmented, flattened, and hallucinatory—assuming that these classes can be clinically distinguished. Nor does this require that any of these be regarded as a subset of another.

In describing the observations in this way we have plotted them according to their proximity to certain criterion events whose projections have been clustered along four parametric reference axes—anxiety, fragmentation, flattening, and hallucination. But the behavior observations we are to classify are not, in this usage, related to the reference axes directly; instead, they are understood in terms of their proximity to clusters of other observations which have been so rated. This is what we do, for example, when we use a ward behavior rating scale and come up with a diagnostic profile of each patient's behavior. This procedure permits some retention of the abstractness of the classification system, although it sacrifices a certain amount of conceptual precision and flexibility. Moreover, one need not find himself hopelessly confronted with the barriers that are erected by disease entity thinking or by other forms of categorization that raise dogmatic obstacles to creative or analytic inquiry.

10. Facilitating the Clinician-Theoretician Exchange

Ideally, from the point of view of a personality theorist, it is most important for any system we use to maintain a clear distinction between events and reference axes—or between observations and constructs—keeping in mind particularly the postulated nature of the latter, and keeping in mind also the interplay between the two in producing meaning for events, and tangibility for constructs. If, for example, a clinician speaks of anxiety, the theoretician will shudder at any implication that anxiety was what was observed, instead of being the postulated dimension of reference for the clinician's observations—or the cluster of behavioral criterion referents he judged to be similar.

The theoretician is always concerned with envisioning fresh ways of looking at man, and if what is brought to his attention comes all sealed up in some clinician's favorite conceptual wrappers with a note that it must be taken or left in the state in which it was delivered, he is likely to feel that there is nothing he can now contribute except, possibly, to bundle up the packages into gross cartons without breaking their seals. This is precisely what happens to the research psychologist when he is asked to "come in and do research" on a 20-year accumulation of clinical records. All he can find in the files are the wrappings within which the original observations were once classified. The observations themselves, which might have provided some grounds for reconstruing patients and their behaviors, have long since been lost—probably never got beyond the interviewing rooms where they originally occurred. Doing research on case files has, in my experience, always proved to be a waste of time as far as personality theory is concerned. In fact I am inclined to argue that the nosological system in current use does more to impede the development of improved personality theories than to ac-

celerate it. If we cannot replace it, we ought, at least, to abandon it.

11. Universal Reference Axes— Now, if Possible

But if the convocation of this conference means we have now reached that stage of scientific and philosophical enlightenment where we can stop erecting categorical barriers in dealing with patients, and can regard the patient as the entity rather than the disease as the entity, and, furthermore, if we are at a point where we can use our constructs as multiple reference axes—plotting each set of observations against a set of key constructs, as physicists do with their notions of mass, time, and energy, then this conference may prove to be most timely.

Perhaps I should have attempted to suggest what such a key set of constructs for psychiatrists might be. I don't know whether "anxiety," for example, could be made a universal reference axis against which psychiatric records would project all diagnosed cases. Certainly it is a notion widely employed in diagnosis, and in spite of the variety of notions about its derivation, seems to lend itself to considerable consensus in clinical observation. The construct of "projection" also suggests itself, though it may have picked up too much of a constellation of other constructs. It therefore might better be replaced by a more propositional reference axis, such as the I-E axis—internal versus external perception of the locus of control. Certainly "schizophrenia" is much too constellated in current usage to serve as a reference axis, but "loosened construction," often used in one form or another as a term in putting together a diagnosis of thought-disordered schizophrenia, might have possibilities.

There are many others that might be candidates for a universal system of reference axes, but I have here limited myself rather strictly to the topic assigned me in this conference—the role and function of

classification in personality theory. Besides, the proposal of a complete set of significant constructs is a rather large undertaking. I doubt that I am up to it.

Even if it does not prove practicable for the participants in this conference to go all the way and recognize the appropriateness of postulated reference axes that give the best theoretical structure to clinical observations, and if we must, instead, temporize by endorsing the notion of a non-categorical classification system based on relating our observations to two or more clusterings of events, we shall have made some worthwhile progress. Simply to look at a bit of behavior, first in terms of its gross similarity to one type of criterion cluster—or set, if you prefer to relate this to mathematical set theory—and then to view it in terms of its similarity to other sets, is at least to put it in a concrete perspective. Certainly a concrete perspective is better than none at all. Each concrete referent set can be used as a vantage point from which to view the behavior of the man, and while we may be pretty fuzzy about what abstract constructs are being invoked, perhaps at long last we can break out of our categories and come to see the man from the perspective provided by simultaneously having more than one viewing angle.

Supplementary Comments by Dr. Kelly as Presented at the Conference

I would assume that all of you who wanted to read the paper have read it, and that if you didn't read it you don't want to have it rammed down your throat.

I suspect that what happened in most cases, if I can use my own behavior as an example, is that you looked at the first two paragraphs on the first page, the middle paragraph on the second page, the footnote

on page 7—except that there was none—and then to see if the last paragraph had more than three sentences in it. And if I tell you any more than what I tried to cram into those particular paragraphs, I should be telling you more than you care to know.

Let me talk a little bit about what I think the function of a theory is. The function of a theory is to go beyond merely the embracing of observations comprehensively and economically and lead us to anticipate events which we have not yet witnessed.

Now a good many people describe "theory" simply in economic terms or in comprehensive terms. In that case what you have got in the theory is an encapsulation or the simulation of a lot of facts. This may be nice for bookkeeping purposes or to keep the filing cabinet from getting cluttered up or to keep the IBM people from buying too many of those cards, but it really doesn't provide the kind of service we need.

As I see it, the purpose of a theory is to extend inquiry into the unknown. Thus a theory should be, I think, first of all fertile, and the questions of economy, parsimony, comprehensiveness, testability, and that sort of thing, while important, become secondary.

Not only that but one of the intentions of a theory is to provide grounds for a second round of observations. Thus a theoretician should provide a basis for the clinician to take a second look at his cases and to see if now he can see things which before had escaped him. One way to say this is to suggest that the purpose of a theory is not just to condense what is observed but to enable the clinician to transcend the obvious.

If you encourage clinicians to make observations for year after year within the same framework, the whole undertaking grinds to a halt. What, instead, should happen, with the proper interplay of clinical observation and theoretical speculation, is that new clinical observations

should continually emerge. I suppose the computer boys will have to unload their reels of tape and start over again every year or two, particularly if the dialogue between the theoretician and the clinician becomes really important.

About the structure of a theory: As I see it, a theory provides the warp in a fabric which may comprise warp and woof, the one being the observations that you make and the other being the constructs that we erect for giving us perspective on those observations. Thus, a theoretical structure is an interweave of events and constructions. The events, of course, supposedly stand on their own, but the constructions are the responsibility of the theoretician. And he cannot blame those constructions on anyone else, or even claim that he discovered them on Mount Sinai or in any other laboratory.

He must take full responsibility for the theoretical constructions which he weaves into the events.

Or if we wish to be more phenomenological in the way we say this, we won't say the "events"; we will say our "observations." But, as of course you know, even observations are biased by implicit theory. And theory, in turn, is palpably limited to those observations which we actually have made.

Thus the constructs—these are our own reference axes. They are not properties of the events. They are what we erect, against which we plot the events. The variety of constructs we devise provides multidimensional depth to our observations.

One observation can come to mean a great deal to a skilled clinician who is theoretically sophisticated. In turn, the observation provides substance to his constructs; otherwise, they become not abstract but merely abstruse. So you get definition by intention in the system by the nature of the constructs. But the constructs get their definition by extension, from the events which they place in array.

Now we have been talking a good deal about computers and I agree that they are extremely important and valuable. Some of the results of computers have been very helpful to me—I think they have. Maybe they've led me down the garden path, but I think they have been helpful. What you do with computers is eliminate the redundancy in this interweave of constructs and observations. In other words, you can eliminate the excess meaning from the events and you can reduce the number of events necessary to provide operational definition for the constructs. Thus you shrink the matrix by use of computers so that you have a minimum number of construct dimensions to provide the kind of depth in which you have been able to observe, and you can minimize the number of observations necessary to establish operational definition of the construct.

But this kind of condensation which appears to tell us a great deal in a few terms is not abstraction. It may be abstraction in a lawyer's sense, but not in the sense of a theoretician. It can, of course, provide some clearing away of the clutter so that the theoretician can think more clearly. It's an application of Ockham's razor. But we need to keep in mind that Ockham's razor is a surgical instrument and not a creative tool. It can perform a useful service in cutting out excess terms, but it doesn't necessarily provide us with the construction for a new interweave or a new fabric.

Now there can be a dialogue between the products of the computer and the clinician. Having cut out the excess terms and reduced the number of observations necessary to establish a theoretical axis of reference, the computer can feed back to the clinician. He then, with this reduced matrix, can with more precision do what he was already doing. This doesn't mean he is any more creative or that he has a new perspective, but he may have greater precision and this alone is quite worthwhile.

As I was saying a while ago, it seems to me that one of the main criteria by which we should judge the value of the theory is its fertility. Another way to put it is to say that a good theory does not necessarily provide the answers to questions, but generates better questions. Only at the risk of triviality does it provide final answers. To a personality theorist, the best result of a particular inquiry is not one menopausal conclusion but at least two pregnant questions.

We have been talking about classification and sometimes have used the term "classification" when we were referring to a particular kind of classification which might better be called categorization. Categorization is a preemptive kind of classification; once put into that pigeon-hole, an observation can't go into any other—as a patient once put in a ward cannot simultaneously be put in another.

Categorization is an invitation to conclusions — menopausal conclusions — and hence to dogmatism. In the eagerness of all of us to have some sort of definite discipline, some of us are inclined to think that we ought to have a science—something with a big dictionary containing 18,000 terms and a great deal of mathematical sophistication. In our eagerness to have that kind of science we are likely to resort to categorization. Then, of course, it is only a very short step to the pure dogma that stands in the way of all creative thinking.

Well, the problem is then: Can we now go directly to parametric reference axes and refer the observations we make in the clinic to these reference axes, or must we do something intermediate between categorization and parametric axis reference? Must we tentatively use noncategorical classification in an abstract way? This would mean that we would set up clusters by some of the devices we now employ and then, having set up the clusters, any observation we made we would relate to one or more clusters, according to its nearness to them.

This is a sort of intermediate stage between dogmatic classification—something a theorist hates because it leads to conclusions and not to questions—and pure abstraction, which is a kind of a rough thing to handle when you have to treat a screaming patient in a hurry.

Abnormal Psychology: Nosology Adrift

Kenneth R. Hammond, Ph. D.¹

Reading the papers prepared for this session quite naturally leads me to begin with a typology. My typology involves three groups: Group I consists of administrators, psychometrician-taxonomists, and chemotherapists. This group is strongly in favor of developing an authoritative, useful nosology, and, of course, it is well represented at this conference. Group II consists of clinicians, mainly psychotherapists. Their attitudes toward typology, taxonomy, and nosology are indicated by the respondents (and nonrespondents) to Dr. Weinstock's questionnaire. They consider all this not only a bore but a menace. This group is probably not so well represented here. Group III in my typology consists of experimental psychologists—the laboratory psychologists who are supposed to be doing the basic research that is supposed to provide groups I and II with the scientific basis for classification and treatment. In effect, their role is to legitimize group I and II. Few, if any, members of this group are present at this conference, because, as we all know, they would have nothing to contribute. One of our firmest traditions is that group III, experimental psychologists, does not interact with group I or II. But, neither do groups I and II interact, although

¹ Institute of Behavioral Science, University of Colorado, Boulder.

the wall may not be so high. The difference may be that groups I and II are slightly embarrassed by their failure to interact, whereas group III takes pride in its isolation. The end result is the same, however. When groups concerned with classification, practice, and research within a discipline do not provide one another with information they can use, the discipline does not advance.

The independence and isolation of these groups stands in the way of this conference. For unless nosology is derived from basic research and unless it bears some functional relation to treatment, it becomes mere record-keeping.

My first argument, then, is that it is no use trying any longer to develop a nosology by clarifying terms, narrowing definitions, and using ever more refined statistical techniques to separate truth from error. It has been done many times, and despite David Shakow's suggestion that one more good try might do it, I see instead that his own work indicates that the good try has already been made. Nosology must grow out of fundamental research on psychological process and treatment, not out of consensus. Is this possible in the near future? I do not think so and in the remainder of this comment I will indicate why I think it is not possible in the short run, and what must be done in the long run.

The prime difficulty is that for the past 50 years the complexity of personality theory has far outrun the capacities of basic researchers to deal with it. Since research could hardly challenge personality theory, there has been hardly any change in it. And since researchers do what they can do, they challenged instead simplistic models rather complex ones. As a result communication between research and practice has never developed.

It is worth while taking a close look at the fundamental concepts of personality theory, for we should be clear about why it defies both experimental and taxonomic research.

Three basic concepts form the core of all personality theory which guides all clinicians: (1) Intersubstitutability, (2) irreducible uncertainty, and (3) linearity and non-linearity. Each of these can be explained briefly because each is familiar to you, although you may use different terms.

First, "intersubstitutability" refers to the fact that impalpable central constructs or concepts express themselves in various palpable behavioral forms which may substitute for one another. No one disputes that, no modern psychologist does without this concept, yet neither psychometrician nor experimentalist has been able to cope with intersubstitutability.

Second, "irreducible uncertainty" refers to the fact that all personality theorists consider the relation between impalpable personality constructs and behavioral product to be a probability relation. This irreducible uncertainty is not due to mere technical difficulties of measurement; the probability is there as a matter of principle. The nature of man is such that an uncertainty relation exists between depth and surface; some degree of uncertainty is irreducible in principle.

Third, every personality theorist assumes that both "linear" and "nonlinear" relations exist between construct and behavior datum, and clinicians make such a point of these relations that no more needs be said about this.

There is a fourth point which must be mentioned: It is a general research principle held by all personality theorists; it can be labeled "nonfragmentation." Personality theorists do not wish to be interested in research unless all three characteristics of human behavior listed above are present in the research. And the only research instrument which can fully satisfy that demand is the human judge, the clinician himself. There is, of course, no solid evidence to support the clinicians' confidence in this instrument; but there is no clear set of evidence against it either as the kind of

instrument the clinician believes himself to be.

Grant that I have correctly identified three concepts and a general principle which form a minimal core essential to all current personality theory which guides all clinical work. Why is it that basic research and taxonomical research cannot cope with these essentials?

Basic researchers cannot cope with them because their design procedures are suited to simplistic theories, not complex ones. Their procedures, in fact, are obsolete. Still struggling to isolate one variable at a time, still arranging experimental conditions to fit a simplistic conception of human behavior that no one outside intramural circles considers interesting or relevant, current experimental design and doctrinaire training preclude the possibility of dealing with high-complexity theory.

Statistical taxonomists and administrators cannot cope because they assume a typology as a point of departure—the very assumption which both basic researcher and clinician reject as a menace. (Indeed, this is one of the very few points on which clinicians and experimentalists can agree.) The administrator's need for a typology is understandable and important. He would like very much to know how many patients are in this category and that one; and he would like very much to know how many patients in this category and that one respond to which kind of therapeutic agent. The statistical taxonomist, of course, rises to the challenge presented by the person who needs the taxonomy because the mathematical tools at his disposal are becoming ever more sophisticated and ever more powerful. Modern techniques should be able to cope with a problem which is of obvious significance, and which cries out for the use of such techniques. Unhappily, what appears to be a challenge to statistical ingenuity is also a problem which is conceptually nonexistent. The taxonomists' ingenuity is left to expend itself on an ever-

increasing elaboration of the obvious—analysis of behaviors which are so extreme as to minimize the properties of personality theory mentioned above. In short, taxonomists are left with the problems which are minimally associated with the personality theory which guides the work of clinicians. Those problems which *are* associated with personality theory are perpetually postponed; they are not dealt with because almost no one will identify problems of personality theory as taxonomic problems. In the end, taxonomic research, even if coupled with successful chemotherapy, is not likely to provide much more satisfaction to the personality-oriented clinician than electroshock therapy did.

Is there no way to develop a nosology based on the experimental study of the psychological processes inherent in high-complexity personality theory? Clinicians (group II) cannot be expected to show leadership and creativity because their problems are too immediate and too demanding, as Dr. Weinstock's experiences show. Nor can the experimental psychologists (group III) be expected to change their ways. They have a system for doing research; they do it, they publish it, they train students to do it the same way, and if no one but other experimentalists shows any interest in it, so much the worse for the outsiders. Because the system is highly viable, I don't see how anyone can expect experimentalists to change in the near future. Unhappily, the members of group I, so well represented here, seem to want to pursue the problem of nosology in old, familiar, and unsuccessful ways also.

If the future does not look bright in the short run, what should be done in the long run? If you agree that nosology should grow out of basic research and practice, then perhaps you will agree that nosology must grow out of experimental research involving high-complexity theory. However, since classical experimental research procedures cannot cope with high-complexity theory, one or

the other must change if we are to accomplish our objective. Which can it be?

Since neither personality theory nor classical experimental research can claim to be self-correcting systems, and since both are firmly rooted, the choice is difficult. Indeed, I can only offer you my own resolution of the difficulty and provide a very general discussion of why I made the choice I did and give an indication of the results of that choice.

I have adopted the research paradigm offered by Egon Brunswik over 20 years ago. I adopted it because it makes it possible to cope with the three basic concepts of personality theory mentioned earlier, as well as the principle of nonfragmentation. Let me give an example of its utility as it applies to the problem of studying the clinician as a diagnostic instrument.

Since the clinician insists on the superiority of the clinician as a diagnostic instrument (the implication is that he can cope with the three basic concepts of personality theory as well as with the principle of nonfragmentation) we have found it useful to study the cognitive processes employed by the clinician in his effort to make a behavioral diagnosis.

Now it is important to note that in order to make a behavioral diagnosis one person must learn to predict another person's response to some feature of the environment. Person A must somehow learn to predict what person B will do under various circumstances. What scientific knowledge do we have about this essential, fundamentally important process? None whatsoever. There is no scientific evidence on the subject. Classical experimental psychology has never (and by virtue of its narrow character, will never) get in contact with this process. The Brunswik research paradigm brought us in contact with it, made it possible for us to do research on it, and made it possible to learn something about it. The reason for this is that the Brunswik paradigm takes intersubstitutability as its starting point and al-

lows it to occur, rather than eliminating it in the interest of control.

The importance of interpersonal learning for psychology in general and for diagnosis in particular should be obvious; let me make it more plain by reference to the second essential feature of high-complexity theory—irreducible uncertainty. This topic has to be discussed in terms of diagnostic devices or tests. First, what is the function of diagnostic tests? Their function is to remove reducible uncertainty; tests should “engineer” uncertainty out of the diagnostic system. The advance of medicine in the last 50 years is illustrated in part by the enormous development of apparatus which makes it less and less necessary for the diagnostician to depend on his uncertainty-geared, intuitive, cognitive processes. In psychiatry, however, there has been no advance worth mentioning with respect to behavioral diagnoses. The human observer is still the final judge—and it is his uncertainty-geared, intuitive, cognitive processes which are the main instruments to be used in learning about another person. Because the Brunswik research paradigm will permit and encourage the analysis of the process of interpersonal learning, it will permit and encourage the development of devices which will engineer as much reducible uncertainty as possible out of the interpersonal learning system. And Brunswik's methodological paradigm will permit and encourage this because it incorporates explicitly uncertainty between surface and depth—something which classical psychology has never been willing or able to do. Having accepted uncertainty, we have developed techniques to deal with it; classical experimental psychology has not. As a result, we can speak, and have spoken, to the problem of diagnosis; classical psychology cannot.

And finally, we come to the problem of configural or nonlinear relations between surface and depth. I have taken the position that the proper research paradigm should permit the simultaneous appearance analy-

sis of both linear and nonlinear relations. We have carried out such research and have learned that in one instance, at least, diagnosticians were indeed responding to nonlinear relations, but were doing so to their disadvantage. A happier result, and far more important result, was obtained in another experiment when it was found that humans could learn to utilize both linear and nonlinear relations simultaneously in a multiple-probability learning task, a finding which, if accurate, would have clear implications for the supposed utility of clinicians as diagnostic instruments.

These examples are intended to do no more than indicate that unorthodox methodology is needed to cope with high-complexity theory, that at least one unorthodox methodology is available, and that unorthodox methodology can bring data to bear on personality theory. Whether such methodology can succeed in challenging and thus advancing personality theory, and whether such advance can produce the nosology that is needed remains to be seen.

There is one thing which can be seen clearly enough now, however, and that is the utter futility of the monotonous, mindless, continuous factor analysis (of any variety) of psychiatrists' ratings of patients. Haven't we already learned over and over what we knew in the first place—that psychiatrists differ from one another in their ratings? And haven't we learned that we are going to learn no more than that? Why should we spend our resources in the lacedged analysis of data produced by, say, three psychiatrists?

Why do such studies continue to involve the energies of persons dedicated to research concerning mental illness? Why do such studies play so large a role in this conference? My answer is that psychiatry has no other data which it finds worthy of analysis. There is, in short, nothing else to be done. As pointed out earlier, there is no discipline which psychiatry acknowledges as basic to its practice. (There is, for example,

no discipline that stands in relation to psychiatry as physiology or biochemistry stands in relation to internal medicine.) Therefore there are no basic data other than those provided by the judgments of psychiatrists.

This situation is particularly unfortunate in a discipline which is failing to solve its problems. Indeed, in a discipline which is failing to solve its problems, which can claim no significant advance in the scientific understanding of its subject matter, it is at least ironic to find that the intuitive judgments of its practitioners form the basic data of its research effort. But if this circumstance is ironic, it is also understandable as long as the fundamental difficulty remains—lack of contact between psychiatry and a basic research discipline. Until that problem is solved the evermore esoteric mathematical analysis of expert judgment will continue—much in the style of the evermore esoteric analyses of the hidden meaning of the words of the authorities of the Middle Ages. In the meantime, our scientific understanding of the behavior of patients fails to increase.

Some Implications of Classification in Sociology for Problems of Classification in Psychiatry and Psychopathology

John A. Clausen, Ph.D.¹

Although this is a "discussion paper," it may be fruitful to begin with brief consideration of a problem in sociological classification. The sociologist is primarily concerned with those aspects and products of human behavior that derive from the fact that men live in groups or societies. The basic problems of sociological classification are not too dissimilar from those

¹ Professor of Sociology, University of California, Berkeley

of psychiatric classification. Both fields are concerned with normative behaviors and with deviation. Both must cope with cultural variation. Sociology is far from being a rigorous science but has produced some knowledge of the characteristics of group life and their consequences for the development and behavior of the individual which has proven useful to psychiatry and psychology.

In the early days of sociology and anthropology, theorists often attempted to characterize whole societies or cultures in terms of stages in a presumed evolutionary process or in terms of global typologies. As scholars began to look more closely at the phenomena of social life, however, they inevitably had to concentrate on particular facets of social organization or particular processes within society. Certain realms of institutionalized behavior were, indeed, so salient that their study had already been staked out by emergent disciplines—economics, political science, education. Within these areas, whether claimed by other disciplines or remaining available for claim by sociologists, features exerting a pervasive influence on human thought and behavior were distinguished.

To take a specific example, it was early evident that in most human societies some individuals or groups had higher prestige than others and that, indeed, a hierarchy of prestige rankings almost always existed (1). Both the bases for allocating prestige and the permanence of rankings tended to vary considerably from one society to another. In caste societies, one's position in the prestige hierarchy is given at birth and tends to be retained throughout life. One's associations and his life career are largely dictated by caste. In so-called open class societies, on the other hand, one's position in the social hierarchy depends both on the status originally provided by one's family and on one's own ability to better his position in competition with others. The sociologist who focuses on social stratification is interested in

establishing the criteria for ascription or achievement of status and the consequences, both for the individual career and for the functioning of the society, of the individual's placement and his status-relevant attributes.

One may wonder about the relevance of these observations to problems of psychiatric classification. There are two aspects of the sociologist's interest in such facets of social organization and stratification which seem to me to relate significantly to the issues that confront us in this conference. First, there is the matter of trying to encompass all of the potentially significant aspects of variation in a single set of global classifications, when one is dealing with such enormously complex sets of interacting systems as societies or persons. Social scientists have, by and large, given up the attempt to establish a classification of societies as wholes. They have rather attempted to characterize subsystems and to examine their interrelationships, their characteristic patterns of operation and, to a lesser but increasing degree, the mechanisms that tend either to promote or to constrain variation or deviation.

Psychiatry is not merely concerned with classification in order to understand and generalize about human behavior. It has operating responsibilities which appear to demand diagnoses or holistic classifications. But as has been clearly apparent in the papers delineating some of these operational needs, no one set of diagnostic categories or other classifications is going to serve them all. Can one delineate subsystems of orientation and behavior which can be reliably classified and which, by specification of their interrelationships, can yield better bases for varied problems of accounting, prediction, and control than the present diagnostic categories?

The sociologist's interest in such phenomena as stratification has a second relevance for psychiatric classification. Drs. Gardner, Gruenberg, Kramer, and Cole have noted in various ways the need to

take into account the cultural background or social matrix from which the patient emerges or to classify patterns of social adjustment in addition to diagnosis as such. We are not merely confronted by diseases but by persons whose biological, psychological, and social nature and functioning are closely interlinked. To increase the sophistication of our research on psychopathology, we shall require the development of concepts and categories which deal with the linkages between biological, psychological, and social systems (2)—for example, between placement in the social organization, characteristic patterns of development, and the organization and content of the characteristic behaviors of the individual and the responses of others to him.

An example from social psychology will illustrate what I mean. The concept of role, which may be defined as the set of duties, responsibilities, and privileges that go with a particular position in the social organization (father, doctor, policeman), has been extremely useful in linking personality with the social structure, although no one has been able to propose an overarching typology or taxonomy of roles (3). Roles do not, of course, determine behavior. A role provides a framework of expectations and demands that impinge on the occupant of a particular position, expectations and demands that are variously interpreted and responded to by the individual incumbent. One may study the conflicting demands that inhere in particular combinations of positions which the individual occupies or the problems deriving from ambiguity in certain types of role definitions, and on the whole it appears that greater understanding and predictability of behavior can be obtained by exerting efforts in such directions than by trying to elaborate a taxonomy of roles at the present time.

The concept of role is also germane because some of the problems dealt with by psychiatry and included under the heading

of psychopathology are essentially problems of role performance rather than problems of illness. The person who repeatedly loses jobs or disrupts organizations because of difficulties in authority relationships, and who may at times seem paranoid in his interpersonal perceptions, is a case in point. It is unlikely that one will ever be able to encompass the range of such problems in role performance under a standard nomenclature. To turn now to closer examination of the ways in which a sociological perspective may be useful to our deliberations, I should like first to describe briefly certain generalizations that derive from sociological research on mental illness and second to state a few of my own convictions regarding points made in certain of the basic program papers.

Sociologists have been involved in the field of psychiatric research and (less often) service in a number of different capacities. Some have directly studied patients or patient-staff interaction in treatment settings but more often sociologists have worked with data generated by clinical services. Insofar as they have participated in epidemiological research, they have tended to base their classifications of patients upon those provided by psychiatric staff and to concentrate their own classification efforts upon the social environment or experiences of the patient—social status, social isolation in childhood, characteristics of the social milieu, family patterns, etc. Occasionally, however, they have taken case records and have attempted to analyze the kinds of behaviors manifested by patients and the kinds of designations of patient behavior made by staff (4). When this has been done, it has become apparent that social background and social stereotypes often influence the categorization of patients in psychiatric services as in other realms of social life.

We may note some of the kinds of generalizations deriving from social science research in the field of psychiatry that have implications for psychiatric classification:

One; Different segments of the community or the population send varying proportions of persons to clinics and mental hospitals (5); the patients come to these services through different routes (police vs. family referral for example) (6); and social characteristics strongly influence the nature and duration of treatment received and other aspects of the patient career (7). Two: The characteristics of institutional arrangements and methods of dealing with patients influence the self-image of the patient and his response to care (8). Three: Features of the sociocultural background in which the individual has developed influence the symptom patterns that are manifested in mental illness (as in the studies of Irish and Italian schizophrenics) (9). Four: There appears to be a relationship between the person's social roles—especially the nature of involvement in social relationships with others—and mental disorders, as indicated in the high rates of symptomatology and of gross impairment as well as the high rates of hospitalization among aliens and other migrants, the alienated, etc. (10).

More generally, both the organization of the personality and the orientation of the person may be expected to show influences of the culture and of the individual's position within the social structure. Research in this area has been voluminous but often unsatisfactory from a methodological standpoint. Many of the very insightful studies in the field of culture and personality have not relied upon separate assessments of cultural configurations and of individual personalities but have inferred both from a single process of conceptual construction, often resting more on art than on rigorous methodology (11). Therefore there has been little attention to the relationship between the varieties of deviant behavior and the normative structure of society. Perhaps the most impressive attempts at systematic study of individual personality in a cultural matrix have come in efforts to understand de-

linquency and drug addiction within the subcultures of the urban slum (12) and in the examination of personality constellations within populations subject to acculturation pressures (13).

To a substantial degree, manifestations of intrapersonal and interpersonal difficulty are likely to depend upon the kinds of demands that are placed on the individual in the course of his development and of his functioning as an adult. The discontinuities between norms relating to behavior patterns in childhood and adolescence and the demands of adult roles would seem somewhat greater for the lower class than for the middle class person in our society (14). Where role discontinuities and role failures are acute, they may be defined in different ways in different environments. The availability of a stable network of supportive friends and relatives who can assist in finding alternative role possibilities constitutes one type of difference. Another is the availability of stereotypic explanations of role failure which permit the individual to disclaim that he has, indeed, failed. To the extent that there are differentials in the availability of such explanations, they will have a bearing on the consequences of role failure for the individual's self-image (15). Such problems are devilishly difficult to study even by longitudinal research. They are more often illustrated than studied systematically, yet greater efforts to trace the consequences of social responses to role failure may help to clarify problems in the classification of outcomes.

A discussant ought at least to indicate what he most strongly agrees with or disagrees with in the papers under discussion. I have already indicated my strong agreement with a point made in somewhat different ways in nearly all the papers: we need multiple approaches to classification to permit a variety of analyses of the relationships between specified manifestations of mental illness and their antecedents, their courses and outcomes under

different intervention efforts. We shall learn almost nothing new about mental illness by revising nomenclature, though revision may be helpful if some categories currently in use are clearly unsupportable by evidence. We shall learn little new in improving the reliability of assessment of identical bodies of anamnestic and clinical data, even though reliable classification is an outcome devoutly to be wished. But we shall learn much that is new if we ask questions that require more adequate observation and data collection (entailing conceptualization and classification) relating to such matters as the processes whereby we deal with persons presenting mental health problems; the nature of the social matrix in which they have been incorporated (or from which they have been extruded); the interactions between the labeling process, the treatment process, and the self-concept of the patient; and the relationships between the decisions made on the basis of initial study of the patient and the treatment accorded him at a later time.

Whether or not most mental illnesses entail biological deficits or disease, they are unlike organic disease in that they call into question relationships with others that are necessary to stabilize and sustain the personality. A person's ties with others are likely to be significant for outcome in all but a few severe, biologically defined diseases. A diagnostic label too often serves as an excuse for not looking at the person and his situation, but this is the fault of the labeler, not the label. The elaboration of categories dealing with the person and his situation might help correct this tendency.

There is, however, a respect in which the giving of a label has implications for mental illness quite different from those entailed in giving a diagnostic label to physical illness. Szasz and Scheff, among others, have been much concerned with the problem of giving an upset or problematic person a social role by calling him mentally

ill (16). Karl Menninger has dealt with the problem in "The Vital Balance" (17). To what extent do the different kinds of intervention undertaken to help the distressed person result in classifications by self or others which subsequently influence the course of the individual's relationships with others? Are there alternative ways of defining the social role of the patient? To study the phenomena in question we shall need to examine the way in which the scope of his environment and the opportunities it affords a patient are changed by various classifications and attendant interventions.

As a sociologist I cannot claim to be an expert on the varieties of mental illness. I share many of the convictions expressed by David Shakow with reference to the range of work necessary to achieve the sets of categories that will meet the several needs of therapists, hospital administrators, biostatisticians, and epidemiologists. What has increasingly impressed me, in the years during which I have been working in this field, is the relative lack of systematic study of patient behavior over time, the meagerness of the information on which most classifications are made and the limited range of questions that have been tackled with intensive effort of the sort that characterized the group at Worcester. To an amazing degree, assumptions rather than observations underlie much of our classification.

Even when there is observation and rating, it too often covers only a thin segment of behavior. A patient presented at staff, even with a history provided by family members, or a person interviewed for an hour or two in an epidemiological survey can seldom be adequately characterized by what is available in this sample of behavior. Dr. Gruenberg has nicely delineated some of the problems that must be confronted merely in securing reliable classifications as to whether a person should be considered a "case" or not. I am less interested in labeling "cases" than in establishing re-

relationships between various aspects of the person's performance in varied settings and roles and the biological, psychological, and social facts of his existence. An adequate nosology must, I believe, wait upon the establishment of much more valid knowledge based on simpler classifications.

REFERENCES

(1) See BERNARD BARBER, "Social Stratification" (New York: Harcourt, Brace & World, 1957).

(2) For a general discussion of system interrelationships, see WILLIAM CAUDILL, "Effects of Social and Cultural Systems in Reactions to Stress" (New York: Social Science Research Council, Pamphlet 14, 1958).

(3) The concept role, its use and attendant issues are treated in NEAL GROSS, WARD S. MASON and ALEXANDER W. McEACHERN, "Explorations in Role Analysis" (New York: Wiley, 1958).

(4) See, for example, ERWIN L. LINN, "Drug Therapy, Mileau Change and Release from a Mental Hospital," *AMA Archieves of Neurology and Psychiatry*, vol. 81 (1959), pp. 685-704. Also relevant is THOMAS SCHEFF, "The Societal Reaction to Deviance: Ascriptive Elements in the Psychiatric Screening of Mental Patients in a Midwestern State," *Social Problems*, vol. 11 (1964), pp. 401-413.

(5) See, for example, ROBERT E. L. FARIS and H. WARREN DUNHAM, "Mental Disorders in Urban Areas" (Chicago: University of Chicago Press, 1939); also AUGUST B. HOLLINGSHEAD and FREDERICK C. REDLICH, "Social Class and Mental Illness" (New York: Wiley, 1958).

(6) HOLLINGSHEAD and REDLICH, *op. cit.*; also LEO SROLE et al, "Mental Health in the Metropolis" (New York: McGraw Hill, 1962).

(7) HOLLINGSHEAD and REDLICH, *op. cit.*; also ERWIN L. LINN, "Patients' Socioeconomic Characteristics and Release From a Mental Hopsital," *American Journal of Sociology*, vol. 45 (1959), pp. 280-286.

(8) See ERVING GOFFMAN, "Asylums" (Garden City, N.Y.: Doubleday, 1961); also WILLIAM CAUDILL, "The Mental Hospital as a Small Society" (Cambridge, Mass.: Harvard University Press, 1958).

(9) JEROME L. SINGER and MARVIN K. OPLER, "Contrasting Patterns of Fantasy and Motility in Irish and Italian Schizophrenics," *Journal of Abnormal and Social Psychology*, vol 53 (1966), pp. 42-47. See also MARVIN K. OPLER, "Culture and Mental Health" (New York: Macmillan, 1959).

(10) See DOROTHEA LEIGHTON et al, "The Character of Danger" (New York: Basic Books, 1964). Also H. WARREN DUNHAM, "Anomie and Mental Disorder" in Marshall B. Clinard, editor, *Anomie*

and Deviant Behavior (New York: Free Press of Glencoe, 1964).

(11) An excellent discussion of theory and research in this field is provided by ALEX INKELES and DANIEL J. LEVINSON, "National Character: The Study of Modal Personality and Sociocultural Systems," in Gardner Lindzey, editor, *Handbook of Social Psychology* (Cambridge, Mass.: Addison-Wesley, 1954).

(12) See, for example, ISIDOR CHEIN et al, "The Road to H" (New York: Basic Books, 1963).

(13) For example, LOUISE SPINDLER and GEORGE SPINDLER, "A Modal Personality Technique in the Study of Menomini Acculturation," in Bert Kaplan, editor, *Studying Personality Cross-Culturally* (New York: Row, Peterson, 1961).

(14) For a general review of some of the relevant research see JOHN A. CLAUSEN and JUDITH R. WILLIAMS, "Sociological Correlates of Child Behavior," in Harold Stevenson, editor, *Child Psychology* (Chicago: National Society for the Study of Education, 1963), pp. 62-107.

(15) See ERVING GOFFMAN, "On Cooling the Mark Out: Some Aspects of Adaptation to Failure," *Psychiatry*, vol. 15, No. 4 (November 1952), pp. 451-463.

(16) THOMAS S. SZASZ, "The Myth of Mental Illness" (New York: Harper, 1961), and THOMAS SCHEFF, "The Role of the Mentally Ill and the Dynamics of Mental Disorder: A Research Framework," *Sociometry*, vol. 26 (1963), pp. 436-453.

(17) KARL MENNINGER, "The Vital Balance: The Life Process in Mental Health and Illness" (New York: Viking Press, 1963).

Discussion of Papers by Brill, Weinstock, Gleser, Overall and Hollister, and Stein and Neulinger in the Light of Research in Human Development¹

*Lois B. Murphy, Ph. D.*²

Introduction

I cannot help wearing two hats at this meeting: I have been asked to represent the

¹ The point of view represented here has grown out of research done with the support of NIMH Grants M680, M4093, PH 43-65-41, along with support from the Pfeiffer Foundation and the Menninger Foundation.

² Research Psychologist, Menninger Foundation, Topeka, Kans.

field of human development in reviewing the papers on classification sent to me. I assume that this implies an emphasis on research needs and criteria of concepts. But as a psychoanalytically oriented child therapist, I am concerned with the question of how the helping person can be assisted to understand the patient in the way most likely to lead to improvement in functioning.

First, I shall make a few comments on the papers; these have to do with—

1. the role of a need for classification in research and in hospital management.

2. the rationale and history of existing classification categories.

3. difficulties encountered in using these such as: (a) The overlapping and fluid nature of categories of mental illness, changing classificatory status of patients, etc.; i.e., the lack of consistent entities aside from clear-cut organic disease or damage; (b) the variation in interpretation and use of major categories, especially schizophrenia; and (c) the harm done to patients by the application of socially derogatory labels.

The papers range from: (a) Strong pleas for adherence to established categories on the grounds of the research and treatment confusion that would result from major changes, to; (b) a strong plea for hospitable receptivity to the possibility of involving new and more useful categories through modern IBM techniques which would accurately disclose empirical clusters suggesting more useful categories.

Our ethical obligation to present and future patients is enormous. Most of us who have treated, observed, or otherwise struggled with mentally ill patients have seen the important effects of the self-fulfilling prophecy: the times when a patient treated in a State hospital as schizophrenic becomes more so (just as a functionally retarded child often becomes far more retarded when cast into a traditional institution for the retarded), while the expectation of progress helps to evoke it.

Implications of Aims and Assumptions Underlying Classificatory Diagnosis

The problem of diagnosis requires a consideration both of the aims and of the assumptions involved in classifying individuals for different purposes. I cannot deal with the latter in detail but shall concern myself primarily with classification in relation to treatment.

One probably erroneous assumption is that all conditions are classifiable into neat or "clean" categories—on the analogy of physical diseases like measles, or a broken limb.

Perhaps the IBM work will help to sort out what conditions can be regarded as "clean" entities. There is certainly much evidence that schizophrenia cannot be so considered. Whether we look at Shakow's report of interindividual and intraindividual variability, or L. Bender's descriptions of the diffuse developmental malfunctioning so frequently reported in children later regarded as schizophrenic, or at the wide differences in patterns of remission in different cases, we are a long way from any "clean" diagnostic category.

The diagnostic problem also needs consideration of differences between the problems of classification of children going through rapid developmental changes, as contrasted with adults who, though still going through sequential life phases, may change more slowly.

The following specific assumptions should be considered:

1. *Classification assumes stability of pattern.*—Relevant facts: "Change" in many basic aspects of development is typical of a significant portion of children; for instance: 20 to 50 percent of different samples of children changed 20 points or more in IQ on tests between the preschool stage and adolescence—still more at younger ages, somewhat less but still substantially at older ages.

Egregious crimes have been perpetuated through relegation of children to institutions where their development is blocked

by the bleak and depriving conditions of the institution—on the basis of IQ's whose stability was taken for granted, but, in view of our present information, in many cases questionable.

Moreover, "body build" and "pattern of growth" also change for certain children; some have "fat stages" (Stolz), some get very thin and long and then fill out. Constitutional body types are not uniformly stable.

It is well documented that problems and symptoms during the early years may be quite acute and still be transient, and it is very hard to evaluate the probable permanence of a symptom pattern in a young child. The transiency or permanence of a symptom is closely related to the purpose it serves: Thus, "withdrawal" may actually be a strategic playing for time to allow for improved orientation, assimilation, affective containment or inner discharge, planning or fantasy rehearsal of an active coping effort.

"Regression" may serve a similar purpose, or an aim of allowing for rest, comfort, or renewal after a strenuous coping effort or otherwise exhausting experience. In our research records we have cases of regression followed by vigorous recovery, not only renewed effort but increased drive for mastery. This has been noted especially when brief traumatic experiences occurred during or just before a developmental phase (such as the "phallic phase" at 3 to 4 years of age, or the pre-pubertal stage).

Temper tantrums, tics, and other tension-discharge patterns occurring during stages of rapid increases of drive energy (phallic phase, puberty) are often transient, if the child has support for other solutions. Mild speech problems, nail biting, and other such symptoms may serve the need for tension discharge in relatively innocuous ways, without interfering grossly with social relationships or strong response to the environment generally.

In all such instances we can recognize the adaptive value of the mechanism or strategy which permits the child to cope with strain and still to manage its major tasks in such a way as to survive in its own social milieu. From this point of view what are often called symptoms have not only current survival value, but often play a role in the child's total strategy for progressive development, and yield to later socially acceptable adaptational patterns. At the same time, it is common at certain stages of physiological upheaval to regress to earlier paths of defense and coping which are then regarded as pathological, without recognition of their possible transitoriness. Thus, we frequently find psychiatrists classifying adolescent disturbances in terms solely of the symptom-pattern without regard to the passing adaptational stress accompanying rapid internal change with its aspects of narcissistic threats, loss of the childhood self, etc., at the same time as abrupt external changes (shift to junior high school, moving, etc.) are occurring.

2. *We also need to consider the assumption that classification can provide or assist a prognosis.*—This is, again, most often true in the case of gross defects, disease, or damage—although even here the compensatory and recuperative powers of a child are often underestimated; in our normal study sample we have a child who is now functioning vigorously who was (a) not expected at first to recover at all from an extremely severe polio as a preschool child; (b) not expected to ever be able to get out of bed.

The capacity for recuperation from emotional disturbance or mental illness with full support is probably as great as the capacity of some children to recuperate from severe physical illness. The prognosis depends not only on the classification of the nature and severity of the illness, but also on many subtle interacting factors of drive, capacity to struggle, motivation, relationships with the

environment, quality and quantity of environmental support, etc.

As an illustration of difficulties in classification and the possibility of destructive consequences of the resulting confusions, we can consider the problems involved in differentiating the current categories of retardation, "autism," "schizophrenia," in early childhood.

Kanner classifies the mildly retarded (feeble-minded) in three categories ("Feeble-mindedness: Absolute, Relative, and Apparent," *Nervous Child*, 1948, 7, 365-397).

1. The absolute feeble-minded, in whom the pathology is irreversible.
2. The relatively feeble-minded, in whom the process can be altered by cultural training.
3. The apparent or pseudofeeble-minded, including "the autistic" and "schizophrenic" child.

Kanner has used the term "early infantile autism" to describe a syndrome which can be recognized in the first or second year of life, characterized by extreme aloneness, and an attempt to preserve sameness in the environment, inadequate language development, inability to relate to people, obsessive addiction to routine, and displeasure when their solitude is interfered with. They communicate by gestures, only to the extent of having simple needs satisfied and have so little desire for verbal exchange that they either remain mute or use language in a mechanical, noncommunicative manner. Their lack of response often leads to the suspicion of deafness or innate feeble-mindedness.

Psychiatrists in many other countries have noted this pattern which is often confused with mental retardation. Kanner, himself, comments that some of them may be brain damaged, and perhaps all of them schizophrenic. But at a descriptive level they are "sui generis." The longitudinal study of 63 of these children for 4 years or more showed 17 of the children as having a good to fair outcome, being able to attend school

at about grade level although they were distinctly deviant in personality. But 46 (73 percent) had a "poor" outcome characterized by retardation and/or psychotic behavior. L. Bender ("Autism in Children With Mental Deficiency," *American Journal of Mental Deficiency*, vol. 63, No. 7, July 1959), comments that the nonspeaking, autistic child under the age of 5 years, fitting Kanner's classification, cannot be differentiated from other mentally retarded or defective children as far as their ability to function intellectually or socially, or as far as their future outcome is concerned. While the children Kanner studied had come from intellectual backgrounds in many instances, Bender, working with a broader group in New York City, observes that many of them come from a background of defective or mediocre intellectual attainment.

She also notes that often these children have some relative or absolute high points in level of functioning, such as spatial orientation, formboard proficiency, or special memory facility. She notes further that most mental defectives do have some special abilities of this sort. At the same time she states "the lowest potentiality may be the significant function in an individual who is failing in the life task of development or social adjustment." (P. 82, *ibid.*)

She notes Bender's statement that "in every institution for mental defectives, autistic children can be found who are classified as idiots." Bender has shown that in some cases of childhood autism, the cause can be traced back to encephalitis while other cases had other infantile infections or anoxia at birth; thus, she agrees implicitly with Kanner's recognition of the probable role of brain damage in some of the autistic children. More recently, Rimland has asserted that all "true" autism has an organic basis.

The important point for us is the great difficulty in differentiating between cases of autism with potentiality for a normal or better level of intellectual development as compared with those who can only re-

main retarded, and also differentiating between autism and retardation due to organic or genetic factors.

We do not have sharply differentiated entities; on the contrary, there are overlapping patterns of developmental failures, in which mental retardation, brain damage with or without mental retardation, autistic personality with or without brain damage, schizophrenia are closely related and hard to distinguish.

Moreover, each of these is also closely related to developmental difficulties resulting from defects such as deafness or blindness present from birth. According to the report of W. R. Keeler, children who become blinded at some time after birth do not show the same syndrome of autism, primitive perceptual and motor behavior, defective communication and thought processes, and mental retardation which can be seen in some children blind from birth. In the case of infants who have retrolental fibroplasia, this damage is probably only a part of a more total brain damage associated with prematurity and postnatal high oxygen tension.

L. Bender argues that since autistic behavior can be seen in brain-damaged, or severely traumatized, or emotionally deprived, or young schizophrenic children, it can be seen as a defense mechanism whereby such children withdraw or protect themselves from the disorganization and anxiety arising from the basic pathology.

Moreover, those who have worked with children have sometimes experienced children diagnosed as "borderline psychotic" or "autistic" or "schizophrenic," responding under therapy within a short time and behaving as a mildly disturbed or "neurotic" child. I use the conventional terms in quotation marks because they seem to be so slippery in view of the child's capacity to change (in many cases). None of the classificatory labels provides an adequate picture of what is going on in the child. This is a process involving interaction of some combination of brain damage, other vulnerabili-

ties, traumatic or overwhelming environmental impacts on the vulnerable organism with its accumulated conflicts and lowered thresholds, and the variously confused and confusing defense and coping patterns which this disturbed infant or child develops in its transactions in a bewildered and frustrated family.

We all assume that an adequate diagnostic understanding requires the data from a group of professional skills including psychiatric and psychological examinations, pediatric records on development, obstetrical records on pregnancy and delivery, and the social worker's detailed records of early development in the family situation, in relation to the mother's own needs, problems, and relation to the child. Yet, when we obtain these data we try to force them into stereotyped categories which blur or distort the process in the child, and prevent understanding of how to help the child.

Without individually differentiated study of the developmental process (not just the apparent outcome), the possibility of identifying, for instance, the proportion of so-called autistic or even mentally retarded children who may have a possibility of more adequate development, will be missed and the children will be condemned to a life of a defective, which may not have been necessary; and simultaneously a life of expensive dependency on institutional care. Thousands of children have been effectively prevented from recapturing an approximation to normal development through these failures to study diagnostically the developmental process which led to disturbance (often potentially reversible).

When we ask what true entities there are we begin with the following:

1. Those due to localizable defect or injury which may be compensated to a smaller or greater degree.
2. Specific diseases involving disintegrative processes.

How many patients do these account for? By contrast, how shall we think of complex

adaptational patterns which are the dynamic and shifting, changeable outcomes of particular combinations of internal physiological conditions, intrapsychic conflicts or frustrations, external pressures, opportunities or threats? Often here it is more important to assess the degree of transitoriness of the factors triggering a dysintegrative reaction than the nature of the reaction. This we do when we recommend support through a presumably temporary though overwhelming stress. Even experienced psychiatrists often fail to assess the role of temporary growth pressures or the type of support that would be relevant.

Implications

The implication of all this is that as therapists we may find it more helpful to have a concise review of the child's functional difficulties; the entire range of organic and constitutional factors contributing to these (including mild sensory or motor deficits, autonomic lability, sensitivities, etc.); environmental factors; dynamic and other aspects of the interaction between the constitutional and the environmental factors, and the intrapsychic residues of these; resources in the child and in the environment which may be expected to contribute to recuperation; the kinds of help the child and his environment (family, school, etc.) are expected to need.

Labeling and classification play into the mechanization of diagnosis, with concomitant ignoring of important if not indeed crucial individual aspects of etiology. It is a truism that different factors may lead to similar outcomes and that similar factors may interact differently to produce different symptoms. Everything that mechanizes the appraisal of the child reduces the responsibility of the clinician to make himself as refined a tool as possible.

Orientation in the Study of Human Development

Study of human development focuses on principles of growth, learning, and the

struggle both for adaptation to the external world, and to manage tensions arising from inner imbalances, fluctuations of energy, or tolerance, and from cumulative conflicts.

We search for continuities in constitutional traits and are propelled by the force of unexpected changes to understand discontinuities, and the interaction of tendencies contributing to continuity with those contributing to change. Outside of grossly defective cases, very few indeed of the factors regarded as constitutional retain an autonomous status which leaves them immune to interactional change.

Methodological problems, erroneous assumptions, narrow conceptual frameworks, have plagued this field and slowed down the progress of understanding. Although for a quarter-century lipservice has been given to the simple principle of interaction of genetic factors and environment in development the pervasiveness of this principle has not been generally recognized. Hence faulty conclusions stemming from one-sided assumptions have led to tragic mistakes in diagnosis and treatment, as mentioned above (p. 177).

The flagrant error in treatment of the retarded was based on the assumption that the mentally retarded could be reliably classified by tests which were asserted to define genetically determined capacity. Effects of atypical environment, especially cognitive and affective deprivation, with idiosyncratic adaptive tendencies in response to understimulation or overstimulation in early childhood; acute pressure of conflicts, or severe traumata caused by sudden separation from the mother in infancy; variabilities in the pace or pattern of mental development related to illness, mild brain damage; and even the deterioration often contributed by institutionalization were ignored under the influence of the rigid genetic assumption, and the classifications based on this. Even psychiatry ignored until recently the psychogenic factors in a certain proportion of functionally

retarded children under the influence of this rigid concept.

Currently there is in some quarters an equally naive assumption that sociocultural factors directly affect development in uniform ways—without regard to individual differences in thresholds for, or susceptibility to influence by such factors, capacity to resist or to overcome them, or the varying buffer or compensatory factors which may alter their effects in some cases.

Both of these errors are to some degree an outcome of a preoccupation with generalizations in terms of so-called tendencies, which often prevents a careful analysis of subgroups with opposite tendencies which may be masked by the mean because their contrasting directions cancel each other out. I refer for instance to those children whose IQ's consistently decrease in contrast to those whose IQ's increase over a period of time. Only when we had 20- to 30-year longitudinal data carefully examined for sequential trends in individuals (as by Nancy Bayley and Jean Macfarlane's staffs) were we in a position to formulate sounder hypotheses regarding the process of mental growth.

And only when we have refinements within sociocultural data permitting analyses of variations between individuals in reaction to defined conditions with differential individual impact will we begin to resolve conflicts between opposing findings in this area. The Gluecks' study of children who did and did not become delinquent though coming from the same kind of slum environment is an example of progress here.

Drillien's study of premature infants showing the increased incidence of problems as socioeconomic level declined and also as maternal care grew less adequate is an illustration of current efforts to deal with interacting and cumulative factors. However, even this is research at an elementary level. Only when we have studied the interaction of genetic and constitutional factors with environmental factors

in the intimate and cumulative interactions of the developing child and family will we begin to have a solid foundation for understanding of the process lying back of transitory versus cumulative pathology, and the possible contribution of different kinds of intervention.

I have begun by discussing errors of the psychologists and sociologists due to rigid and erroneous assumptions. The assumption of incurability of so-called schizophrenia is an old example in psychiatry; and moreover wherever we are concerned with organic contributions to disturbed functioning we find some professional people who assume a static condition which is irreversible and cannot be compensated. When this assumption is made we find it easy to overlook contradictory evidence.

Our understanding of individual differences in developmental patterns, and of factors contributing to these is still immature. Neuropsychological research, research in perception, and in the interactions between autonomic reactivity and cognitive functioning are all in a phase of rapid development which is changing our view of etiology, of prognosis, and of needed therapy. As these and other sciences progress, classification systems will have to take account of their findings. Lines of research which will influence the assumptions underlying classification systems include study of:

1. The capacity to maintain homeostatic balance.
2. Resilience or the capacity to recover from states of disorganization or dyscontrol.
3. Ego cohesion, or the capacity to maintain coherent cognitive functioning under inner or outer stress.
4. The capacity to seal over traumatic disruptions of integration and to develop stable patterns of functioning "on top of" such layers.

5. The capacity to use restitution for early environmental deficits which led to inadequacies of development.
6. The contributions of these to the level and stability of later development and mental health.

Each of these, or all of them, require intensive psychodynamically oriented psychiatric studies of infants and young children, with followups through successive stages and vicissitudes of development, to see how these patterns are affected by developmental shifts in drives and in level of integration of the organism, and by sequelae of illness, traumata, chronic stress, etc.

It is only recently that fairly adequate tests of hearing and vision for infants and young children permit the assessment of possible contributions of sensory deficit to retardation, withdrawal, and maladaptive coping patterns leading to mental illness. Still less have we studied the extreme sensory sensitivities and affective reactivity of infancy and childhood which contribute to adaptational difficulties; or the early acquisition of a predisposition to anxiety as discussed by Greenacre; or the contributions of phase-specific needs which interact with such sensitivities or predisposition to anxiety to contribute to functionally autonomous tendencies which resist normal recovery tendencies in later childhood. In emphasizing the role of early sensitivities we are, of course, implying not only the possibility but the necessity of considering the interaction of economic and dynamic factors.

Infancy and early childhood, factors in fixation of overdetermined symptoms arising in a context of multiple traumatization during the vulnerable second year of life have not yet been integrated into a comprehensive picture of patterns of early disturbance. The interaction of such disturbed residua of the first 3 years with the conflict-provoking experiences of the often highly emotional period of later preschool years

leaves further patterns to which the adolescent (or adult) may regress under intense or cumulative stress. We do not yet have the kind of long-term developmental studies which would provide definitive evidence as to factors determining the later choice of symptom pattern out of such overlaid residua.

The question as to whether, for instance, the symptoms fit more into a "neurotic" pattern of exacerbation of anxiety, or a "schizophrenic" distortion of thought processes may have to do with other factors which influence the tolerance of acute anxiety. This, in turn, may be related to early constitutional factors.

The Developmental Approach in Child Psychiatry

In the field of child psychiatry Dr. Tarlton Morrow has formulated the developmental problem as follows.³

(1) Normally, the intrapsychic arrangements in children vary enormously over time with development and maturation in some kind of an environment. It has been found desirable—actually after many centuries of confusion—to divide this unfolding process into various stages, or sequential levels, of development. These stages are characterized as "normal" and, although overlapping, usually are considered or described as reasonably discreet—at least for purposes of study and understanding. It is not clear when any of the orders of dysfunction might be actually quite normal for a certain stage of development, nor which and when certain dysfunctions would become abnormal. What would be a coping device for everyday living for a 6-month-old child, for example?—hallucination of the feeding mother?—extensive projection of the infant's oral aggressive urges? It would probably not be too difficult to suggest a hierarchy of disorders for this aged child, but would it apply to the

³ Morrow, J. Tarlton: "Thoughts About Diagnosis and Classification of Disturbances of Childhood," unpublished.

18-month-old infant?—or to the 2-year-old child?—and then again to the 4-year-old child? Since the child is constantly changing, a description to characterize 1 month, or year, of his life no longer applies at some later period. Certainly a rather complicated series of hierarchies, each applying more or less to one discreet age group, or to a stage of development, would be necessary. Then, were these to have been devised, allowances would have to be made for the overlaps between stages, as well as for the individual differences in rate of development of certain children who deviate from the norm (and pattern of development).

(2) Another implication, which is often overlooked, is that part of the task of the ego for children and adolescents is to master the effects of this growth. Each task as it makes its appearance is unique in the life of that individual; the demands made upon the adaptive capacity of the ego are qualitatively different from any demands placed upon an adult. The 2-year-old, for example, is not only coping with the usual instinctual drives and the problems of developing within his mind a workable self-representation, as well as representations of the parents as intrapsychic percepts. He is also actively increasing his ability to talk. This increased use of verbal symbols with all its consequent communicative implications though in one sense a product of the ego—an autonomous one according to Hartmann et al.—also is in another sense a demand placed upon the ego. Never before has he had such an ability and never again will he be asked to deal with such a situation in quite this way. He must at this time of life also cope with the newfound ability to move about much more freely—walk, run (away), jump, climb—as implied in Erikson's use of the term "autonomy" to describes a central aspect of this phase of development.⁴

⁴ A practicing child psychiatrist, when he is given the task of understanding a two-year-old, must therefore take this additional aspect into consideration in evaluating the child.

Therefore, in devising a hierarchy of disorders the way in which the child is coping with his developmental and maturational tasks would have to be taken into consideration.

I shall not deal here with details of differences between the developmental problems of children and adults; I want to emphasize, however, that in adulthood, the developmental tasks, stimulation, and stress or crises are as vivid, and involve as many risks and also potentialities as the crises of childhood: this would seem obvious in view of the breakdowns of 25 percent of marriages, work failures, identity problems of both young adults and retiring senior citizens. These adult developmental stages both evoke preceding adaptational crises and add new factors of organic and social change, which may increase vulnerabilities.

The Need for a Broader Approach to Development

While recent approaches to the role of developmental phases have been helpful, the problem of behavioral and adaptational changes, as we see it through studies of the development of individuals, has hardly begun to be conceptualized as a whole, outside of studies of learning which are often focused on one mechanism while ignoring all others, and on statistical evidence of a single tendency, at the expense of all other interacting tendencies.

Moreover, the new and important neurophysiological understanding of the working of the reticular system, of interaction of different cortical and subcortical systems along with the autonomic and hormone systems, the maintaining of these patterns of interaction, or institution of changes has not been integrated into developmental studies (or for that matter, other approaches to behavior).

I believe such studies will be important for the understanding of predispositions to what are now called schizophrenic, neu-

rotic, and other types of pathology, and the basic questions of reversibility of what are commonly regarded as nonadaptive patterns (although at their beginnings, they may be expressions of the best coping efforts the child could make).

We have not even had adequate incorporation of biochemical and metabolic studies of individual differences including individual differences in variability in these aspects of functioning—nor for that matter, of most of the other zones of individuality as described, for instance, by Roger Williams. When we do, we may be able to predict better which children under given external stress at a vulnerable stage as in adolescence will be likely to react with so-called schizophrenic or so-called hysterical patterns and which ones will recover from this disintegrative phase in response to what kind of help. (I am assuming that as we know more about all these different aspects of the organism and its developmental vicissitudes through different experiences in the environment, we may find less vague, less equivocal, loose, imprecise terms more helpful.)

Interactions between physiological and psychological variables are seen in Witkin's report (personal) that "field dependent" (less differentiated from the environment) individuals are most autonomically labile, more easily flooded, less physiologically (and I would assume, neuropsychologically) differentiated. A related finding is seen in relationships between aggression and certain blood-pressure patterns (in our data and the more extensive data of Fels).

Such psychoanalytic orientations as that of Greenacre on the one hand, and Laretta Bender on the other, will gain solidity from future studies along these lines, I believe. That is, Greenacre's persistent predisposition to anxiety which underlies irreducible neurosis in her opinion, may be seen as a resultant of early infantile flooding of cortical and subcortical systems with biochemical products of frequent autonomic upheavals. And Ben-

der's "massive difficulties in early integration" as this interferes with adequate perceptual differentiation in the infantile stages when cognitive functions are becoming autonomous, and thus lays foundations for so-called psychotic functioning, will also be accessible to precise study in the early months of infancy.

We can, then, have improved predictions of the kind of distortions of development, the types of adaptational difficulties and solutions which children with different configurations of somatic-psychological problems in infancy will show at later stages.

I have given these illustrations because I think they fit with the hope of empirical methods for developing new and more precise ways of organizing information regarding disturbed functioning, expressed in several papers.

Since new methods are generously represented in papers by psychologists in this group, I shall confine myself largely to orientations growing out of our longitudinal studies of development, as these influence our approach to problems of classification.

Like Jean Macfarlane and others studying development from infancy through childhood, we have seen startling shifts from one adaptational level or pattern to another: Unexpected losses or gains in level of mental functioning; dramatic shifts in social orientation as when a previously somewhat isolated or withdrawn child becomes a leader or the most popular in the group; and even changes in body-build as when the slender child becomes a "fatty" and vice versa.

In other words, from a research point of view in longitudinal studies of development, we cannot afford static classifications: When we classify at all, it must be with full recognition that the classification may change the next time around, and our responsibility is to keep constantly alert and receptive to change initiated by growth, by new emergent capacities, by transforma-

tions and reversals, by responses to new opportunities, by solutions of previously ignored or misdiagnosed health problems as well as recovery from well-diagnosed and treated illnesses, from correction of defects; and from a multitude of other interacting factors in the child and in the environment, including progress in mastery of the environment and in resolution of inner struggles.

If the changes occur in frequent swings, we may consider the child generally labile; if they occur in response to significant major changes within himself or in the environment or his relation to it, we may see him as effectively flexible.

Thus, the lability underlying the current symptom-pattern or adaptational style is at least as important as the symptom pattern itself. (Such evidences of previous lability might be a sounder basis for classification into hospital management groups than the symptom pattern.)

Some problems in classification of patients are related to this matter of lability or fixity. How do we judge the degree of internalization or rigidity of anxiety or anger in a person who has been from the beginning and still is, in a frustrating, depriving, or otherwise threatening situation, due either to his own deficits, or to persistent inadequacies in the environment, or to interactions between these? The rigidity of a given pattern can only be judged after behavior is tested under conditions which optimally meet the child's needs. Only if we see him at his best and most comfortable, can we judge which patterns are irreversibly ingrained.

In this connection, it is important to note the problems inherent in the nature of the data which we obtain in an examining situation. If the child is threatened at home, the examination—in a strange place, with strange persons doing strange things or asking strange questions—is also threatening to some children, especially the sensitive ones who from infancy become anxious in strange situations.

Also, the child may be quite different in responding to different sorts of examination, different materials, examiners of different sex, age, or type of approach or degree of situation. Cues from such observed differences can provide evidence in regard to variability of the child in response to different patterns of demand, pressure, or other stimulation.

In view of the fact that reactions to defects, brain damage, emotional deprivation, or sudden traumatic loss (and to other conditions) may, in different children, look like or later take forms which we call retardation, schizophrenia, neurosis, or other pattern, it might be worthwhile to focus more of our diagnostic efforts on etiology in relation to the nature of the individual organism and psyche, and its early experience in and reactions to the environment. One approach to this could be the organized, integrated effort to obtain a clear picture of the basic vulnerabilities and resources of the child, and the ways in which these had interacted and had been modified through the course of development. This review would be concerned as much with the degree and pattern of variability, the response to changing qualities and kinds of stimulation, the areas of persistence and of change, as with the presenting picture.

Discussion of Papers 2, 4, 15, 20, 22, 26

I am not qualified to discuss details of statistical methods and formulae as developed by Gleser, Overall and Hollister, and Stein and Neulinger, although I am in sympathy with the efforts, insofar as their results are subjected to careful evaluation of the consequences through time. A major but not necessary limitation of most quantitative studies at present is that they deal with static data obtained from observations or other reports on individuals studied in one time, place, and adaptational setting, without firsthand data on intraindividual variations from setting to setting, examiner

to examiner, life phase to phase. Thus, statistically coherent outcomes can be utterly misleading from the point of view of understanding long-term interactions between factors contributing to continuity and those contributing to change; or understanding the role of the internal and external conditions contributing to the pattern as seen at one point in time. A number of the papers for this conference do recognize the contingent, often temporary, or fluid character of the behavior pattern sampled at one point in time, one developmental phase. Whether we are dealing with constitutional bodybuild types or personality types, as soon as we give adequate attention to changes through time, our types get more and more complicated and elusive.

In this connection, Brill's discussion of the difficulties in use of diagnostic categories along with the necessity for diagnosis and classification seems incontrovertible; I can only repeat what others in the present group of papers have also noted, that initial diagnosis and classifications often have to be changed after treatment is underway, and that the limitations in present systems are constantly documented by disagreements between diagnosticians. Weinstock provides ample detailed and challenging data on these difficulties which I shall not repeat here, but which are related to my opening pages.

Weinstock's discussion of difficulties of compressing a diagnosis into the expanded APA diagnostic system fits with my own limited experience in using this, for reasons that must be clear also from my previous statements. I would call attention, however, to the problems implied in Weinstock's observations that an increasing number of analysts are questioning the diagnosis of neurosis and suggesting that all symptomatic disorders are cases of schizophrenia. Why use that term with such a pejorative history? It is implied that disorders of cognitive, affective, somatic, and/or behavioral functioning are involved; so why not define

these as aspects of forms of underlying failures in integration revealed after the breakdown of masking control patterns?

I shall discuss Grinker's paper in some detail because I can empathize with the topic, the problems seen by him, the methods, and methodological difficulties which emerged. Many important points here (as in other papers as well) are hurried over in the interest of covering much ground fast. For instance, he noted that "feelings and concerns factors did not correlate at all with current behavior factors, and that apparently the outward signs were not really symptomatic of the depression itself, but rather that they were expressions in a characteristic personal behavioral pattern of an inner distress."

To my mind this certainly implies the need for more careful distinctions between these levels, and more adequate descriptions of the personal behavioral pattern which I would see as a composite of the individual's disintegrative patterns and coping efforts.

Dr. Grinker's implicit challenge of psychiatric stereotypes, and of the current derogation of careful description of the patient could be extended still further. Not only has the natural history of psychiatric diseases been neglected as he correctly states, but the natural history of the patient's development as a whole, within which the illness developed, has been more seriously neglected.

My reflections on this point grow out of data from our longitudinal study: since Dr. Grinker discussed variations in depression I shall deal briefly with this. At the time of the Kennedy assassination the reactions of several pubertal children whom we had known from infancy led us to consider nuances of grief and depression as seen through various aspects of our data.

Janice: "Just felt the United State was coming to its end. I felt I wanted to kill the person who had done it. * * * I also wanted to die. I felt there was no reason for living. It seemed to me that life was over for everything."

Helen: "I was stunned and shocked. * * * I just sat motionless, dazed. * * * I watched TV about 44 hours (preoccupied). * * * I felt as if I wanted to take Caroline, John, and Jackie into my arms to comfort them."

Daryl felt all sick inside at the first news, then she was filled with sadness (sic) and disgust. "I hated the man who did it. * * * my heart was filled with hatred. * * *"

Each of these children had a developmental history which led to depressive tendencies which were however differently toned, and differently managed: Janice, a puny, unplanned baby, runt of the sibling group of five and an unwelcome addition to the family of a menopausal mother, had to fend for herself, amidst the teasing of her brother and annoyance of other family members, especially her mother. With buck teeth, thin hair, and little to compete with two older dazzling sisters it is hard to see how she developed the attractive, competent adolescent personality we saw. At the same time, we were not prepared for the undercurrent of despair released by the nation's tragedy. Janice's depressive tendency seems to be tinged with deep narcissistic threat and a feeling of despair probably augmented by the guilty hostility she feels toward her rejecting mother.

Helen illustrated a predisposition to depression with a different quality. She was also an undernourished and odd-looking child, born to a mother who had lost her previous baby and may have felt that this little stranger was an unpromising addition to the family which included two attractive, healthy, preschool children. With strong oral drives and needs for contact, combined with sensitivity to overstimulation, Helen at 20 weeks seemed to feel deprived by her already again pregnant mother who, though responsible and competent, gave her less intimate attention than the baby seemed to want. With persistent upper respiratory infections, she continued to be undernourished; her craving for contact was expressed in extreme provocativeness and physical exploration of adults as a pre-

school child. Later, her health improved with medical help and her behavior became more socialized, but with a wistful undercurrent of longing for intimacy and tender contact. Helen's depressive tendency seems to have anaclitic roots, and to be relatively free of the hostile entanglement we can see in Janice.

Still another potential depressive pattern can be seen in Daryl, a premature delicate baby whose mother also became pregnant again when she was 2 months old. She was the most hostile child of our intensive group, the most acutely anxious, and has had periods of the most severe withdrawal, with extreme fantasies expressed to her mother at the age of 7, "How could we burn and burn and burn in Hell forever?"

By contrast with all of these, a fourth example of precursor of depression is to be seen in Chester: His infancy was idyllic, and only at the birth of a little brother when Chester was 2 years old did his anger get aroused, followed by stoic resignation as his formerly attentive mother became preoccupied with a new baby girl each successive year for 4 years. Chester's depressive mood is toned by disappointment that early paradise did not last long—a very different quality from the depression felt by the children who never did have what they longed for.

In addition to these variously different narcissistic, anaclitic, hostile-demanding patterns established so early, each child was a different organism with different coping resources, different factors contributing to the stability of the child's equilibrium, different temperaments. Each of these children with varying success tried to solve the problem of disappointment: Helen by religious dedication and helping others; Janice by reaching out for gratification in all available directions and enhancing her appearance; Daryl by engulfing dominating dependency on her younger sister and her mother; Chester by discouraged compliance and sticky oedipal involvement with his fickle mother.

Many factors, including biochemical, would need to be studied in order to understand how drugs could alleviate a severe depression in any of these persons as adults—and which drugs. Their individuality might well be found to be as clear in the biochemical area as in the quality and origins of depression, and the individuality of their coping methods.

If therapeutic work assignments were made, other considerations along with the depression pattern would need to be considered: Chester's adolescent pleasure in construction and invention, Janice's aesthetic taste, Helen's pleasure in animals or living things to be cared for, etc.

Granted that I can present here only fragmentary evidence, it is in each case a relevant piece of a psychological pattern which, though having depressive aspects in common, has a different history in each case; I believe that in all of these children, if mental illness occurs at a later stage, it will take the form of some kind of depression.

There are certain similarities between the patterns seen here and the factor patterns (p. 15) defined in Grinker's analysis: Pattern A was distinguished by apathy, slowed speech, cognitive disturbances, and slight guilt feelings (in a primitive pattern probably reflecting early infantile distress which if more extreme might have been marasmic). Pattern B was agitated, clinging, with considerable guilt feelings, and I would guess was characteristic of patients who were disturbed during the later aggressive phase of infancy. Pattern C was agitated, demanding, and hypochondriacal, with feelings of abandonment, and probably reflects a regression to separation experiences of later infancy. The patterns arrived at by this statistical method would, I think, be sharpened if compared with relevant cases whose etiology had been thoroughly studied. In any case, the therapeutic approach needed is probably quite different in these different patterns.

Grinker also implies that the medical model for diagnostic typology is valid for mental illness, and from this moves to the assumption that empirically derived clinical types should furnish the basis for valid correlations with other systems as well as facilitate etiological and prognostic statements. However, insofar as these depend solely on behavior at one point in time, or at one stage in development, this effort is subject to all the problems I have mentioned.

From the examples I have given, I hope I have documented the point that we need not only a description of current behavior of the patient, and of precipitating events; we also need a full account of the early constitutional including genetic factors, their interaction with the environment, residual conflicts, and the primary adaptational style evolved during the phase of infantile structuring of personality, modifications by subsequent ongoing, and also traumatic experiences including methods of defense and coping, and the resultant layering and transformation of the personality. This can give us a background for understanding the early vulnerabilities and strengths, their contribution to the impact of the environment, the style of reaction to and defense against stress, and the early and later acquired sources of resilience, and the meaning of the behavior as seen in cross section at the time of breakdown.

I would agree, however, with Grinker as well as the others who have been using IBM methods for handling larger quantities of data, that methods could be programed for handling the complex interactions to which this sort of study leads.

A point of semantic clarification: In his paper, Dr. Grinker describes the factors discovered from analysis of patients' feelings and concerns. Factor I is a self-concept of "badness," with characteristics of hopelessness, helplessness, failure, sadness, unworthiness, internal suffering, and guilt.

From a developmental point of view, it may be useful to distinguish "feeling bad" in the sense of hopelessness, helplessness, and suffering, from "unworthiness and guilt." The former is a very primitive feeling—our adult conceptualization of the inarticulate suffering of the poorly functioning infant who experiences constant distress and discomfort. Guilt and unworthiness is a later development after greater differentiation of self, sense of autonomy, internalized aggression, etc., have evolved. Guilt feelings may be, thus, more capable of resolution in response to therapy than the primitive sense of "feeling bad and there's nothing that can relieve it" of the earliest weeks.

Such distinctions as this can be made both on the basis of observations, and of inferences, and of theoretical deductions, when we utilize the contribution of a developmental approach. The fact that in an adult, "feeling bad" and "guilt" are often coexistent does not preclude the importance of distinguishing between these.

(Concretely, I am hopeful of the probable results of child therapy with a child whose guilty feelings developed after the earliest structuring of the personality, and I do not feel that I could hope for basic change in a latency child with a severe feeling-bad orientation which had been chronic from the first six months of life.)

His statement that older people are "more prone to show the aggravated outward signs of depression" implies also the necessity to assess the organic factors of every sort which contribute to disintegrative tendencies and difficulties in controlling, mastering, or compensating for stress-reactions. He does refer to the relation between anxiety and elevated adrenocortico steroids, but a full picture of organic factors will ultimately include far more understanding of interrelations between hormones, autonomic patterns, subcortical and cortical functioning.

Grinker's sensitivity to the range of variation within groups, the wide variability

with respect to anxiety, for instance, again points to the need for fuller understanding of structural as well as phase-specific dynamic and economic aspects of the development of the depressive pattern.

His comment that "most people prove to be atypical"—that is, they do not fit neatly into any system of types and there is far more overlap than separation among groups—again underscores the need for multi-dimensional thinking about individuals. His suggestion that focus on salient attributes can help to sort out a plethora of data can be useful if we have adequate criteria of salience. Without being able to offer you a firm method for doing this I would like to suggest that the variable or group of variables which has most influenced other traits is "salient." Thus, a variable may be conspicuous, but not salient, if it is a visible deviant characteristic but is isolated. Marked autonomic lability, however, may often be very salient, contributing to disruptive internal reactions and external manifestations which interfere with coping and lead to feelings of inadequacy. Certainly much more study of salience in the total spectrum of vulnerabilities and strengths, and both disintegrative and resilient patterns is needed.

Incidentally, this approach would be increasingly effective as hospital records on newborns, and pediatricians' records on the first months of life become more adequate in describing the problems and patterns of adaptation of the young infant, and the growing child.

Summary

The papers available to date (Oct. 1, 1965) vividly document the utterly different situation of the psychiatrist concerned with classification today as compared with Kraepelin's time. Currently there are available resources which that great innovator never dreamed of: The capacity of IBM machines to sort out empirical clusters,

testing old patterns of symptoms and offering new ones; the contributions of psychoanalysis and human development to a flexible, genetically oriented recognition of the interrelations of continuity and change in affective and cognitive functioning. Also, as part of these: New awareness of both the dramatic roles of deprivation and overstimulation or stress on the one hand and multiple subtle and severe vulnerabilities on the other, along with other positive and negative environmental, organic, and intrapsychic interactions in producing disequilibria of functioning which may be expressed through socially disturbing behavior symptoms. Moreover, neurophysiology and neuropsychology with their new understandings of the mutual contributions of structure and function are providing us with insight into problems of arousal, activation, control and integration which might well make a major contribution to a new nosology. Electronic developments even more important than IBM work are making it possible to pinpoint minute structural zones in the C.N.S. (including subcortical systems) where damage, overloading, biochemical disturbances or other interference may contribute to severe cognitive, affective and/or behavioral distortions. Parallel with these are the contributions of drug research which can demonstrate empirically the possibility of changes and control over specific combinations of symptoms—probably implying a capacity to reach such locally crucial sources of disintegrative reaction.

The impact of new developments in therapy whether psychoanalytically oriented (to change in the affect, conflict, or anxiety load) or drug-oriented (to changes in thresholds) or to environmental approaches (therapeutic milieu) is to increase the range of reversibility of symptoms, thus placing the emphasis more and more on a potentially controllable illness process in contrast to concepts of irreversible disease entities.

Supplementary Comments by Dr. Murphy as Presented at the Conference

Hooking onto some of the things that have just been heard from Dr. Clausen's paper, I would like to extend the question of cultural stereotypes and connect it up with a question of international epidemiology which was raised a little earlier, and indicate that it seems to me that we need to know first of all what a culture regards as pathological. That is, those of you who have been around have probably seen in India, for instance, a Sadhu walking around with a painted face and nude, isolated largely from the rest of the world. With us, he would probably be considered schizophrenic. India makes a niche for him. He is accepted. He is part of the landscape. Everybody knows, "Yes, he's a Sadhu." He is all right, and nobody puts him into a mental hospital.

This is one fairly extreme example, perhaps; from the point of view of our situation here on this side of the ocean I think that we might do well to listen to the comments of people from other cultures about us. Here I would like to quote the statement of an Indian friend of mine who came to the United States and who commented that he felt that the demands that were made on young people here to get married, get a job, decide a career, support themselves independently, and so forth, within a short number of years after adolescence, were enough to create a great deal of pathology.

And I think that if he had been reading the papers in the last few months and come across the arguments between Charles Chaplin and his son, he would have felt that this kind of fight might not have been necessary in India because people, at least in many of the traditional areas, simply don't expect the young to make so many adjustments so fast.

In other words, I think that it is quite unsound to assume that particular categories which have their roots in our individualistic and competitive tradition would apply to all cultures around the world. In other words, we have to think relativistically.

Now in relation to what might be a universal type of finding to be searched for, I would like to mention the kind of thing that Pasamanick has been turning up, along with related data and observations and concepts such as Greenacre's "predisposition to anxiety."

I think it would be possible to get an epidemiological study of vulnerability in infancy and early childhood which would be important to have in relation to difficulties of integration and adaptation to whatever the demands of the particular culture were.

Furthermore, it would be worthwhile to look at range of severity of typical expectable stress in each culture in the relation to the supports, and the intrapsychic and behavior outcomes. All of this could be preliminary, before we began to sort out what can be regarded as universal and what is culturally relativistic.

As to the comment that in child psychology there is much that we don't know: if this is true it means that there is much in psychiatry that we don't know. Children become adults—that is, most do; some children don't because we didn't understand marasmus in infancy and other severe childhood difficulties which prevent children from growing up. Here I would include, for instance, certain kinds of extreme psychogenic retardation which we are only beginning to understand.

In relation to the general problems on which many people have commented, I would like to reinforce and perhaps state a little bit differently a problem that many of us have had strong feelings about; that is: the major difficulty with classification and categorization is that it leads to rigid-

ity, to reification, to a kind of nominal realism. This is the problem that was stated earlier: "Give it a name and then we know," instead of concentrating on the question, "What do we really know?"

Hughlings Jackson, for instance, was quite right: even after his complaint about the use of the term "epilepsy," it took nearly a hundred years before we really began to understand it and during that time epileptics were being rejected and prevented from having a decent kind of life.

I would certainly go along with the emphasis on the need for more observations, more differentiated observation, more clarification of the kinds of units and kinds of categories that are meaningful.

In relation to Dr. Clausen's comment that any research on socioeconomic groups, cultural groups, or any other social divisions can provide us with data on certain proportions of pathology, certain kinds of difficulties that come in one group as compared with another: there are, however, certain people who do not show this pathology. In other words, the individuality and the vulnerability of the individual to the particular impacts of the social context also has to be seen, and these have to be seen in relation to each other.

The question of how we can diagnose or predict reversibility is at least as important as the question of diagnosis from either the point of view of behavior as it is, symptoms, or the point of view of etiology.

I have no question at all that we urgently need more accurate fresh description of patient behavior, psychological dynamics, physiological functioning and history, with IBM analysis to promote study of old and new interrelationships of different aspects of functioning at different levels. Such analysis would, I believe, provide us with a map which would show the relatively independent, neighboring and overlapping entities.

Instead of the confusingly fluid shifts of many patients from so-called psychotic or borderline to neurotic categories, or in the case of young children, the discouraging arguments in regard to autism versus schizophrenia, and so forth, we might have a more solid framework rooted in a picture of constitutional givens, with the vulnerabilities and strengths, the early vicissitudes and developmental stresses, in the context of the interaction with the environment, especially the family, and the cumulative conflicts and also strengths resulting from these; current adaptational efforts including both constrictive and useful self-protective defenses, and successful and unsuccessful active coping devices; areas of rigidity and of flexibility, current weaknesses and resources for resilience.

Such data would help us to distinguish between that so-called psychotic who because of gross somatic and psychic disturbance in the early months of life never developed adequate ego-integration, and the one who had significant zones of functional adequacy with other zones of impairment or confusion or vulnerability to specific kinds of environmental pressure or deprivation. The latter will point to the needs which if met can help the patient to recover and maintain a workable equilibrium.

It is easy to imagine that an approach could emerge which integrated these 20th century contributions, viewing the mentally ill patient as a person in whom certain disturbing processes were at work, in which interactions of (1) a pattern of vulnerabilities in the organism—some of which might involve disturbances or defects in the central nervous system or autonomic nervous system, with (2) the patterns of environmental stress—together with (3) the intrapsychic reactions to both—were interfering with the coping resources of the person.

Genetically and constitutionally economic factors would be seen in relation to successive and cumulative phase-specific dynamic

factors, and the structures evolved from the interaction of these with the environment at successive stages. Individuality in each zone, and in both the vulnerable and flexible constructive aspects of variability in the individual would be studied as part of the total picture.

Classifications could include reference to the characteristics of the original organism, positive and negative predisposing factors affecting anxiety level, conflict, disintegrative tendencies; the degree of ego autonomy initially achieved; the developmental phase factors in traumata; the zones of disintegrative reaction; the relationships in the environment as affected by all of the preceding, and the persistent demands involved in maintaining the self-image.

Socially maladaptive defenses and coping maneuvers would then be studied in terms of the multiple kinds of help which could be mobilized, in an integrated effort to improve the person's coping efforts and thus change the current unacceptable adaptation.

And I might say that it seems to me that the more we are able to think in terms of multilevel dynamic developmental etiology—the more we will be able to think in terms of complex mobilization of therapeutic efforts relevant to the needs of the patient.

Problems in Psychiatric Nosology From the Viewpoint of the Biological Sciences

*Seymour S. Kety, M.D.*¹

It may be of some value to hear from a representative of a discipline which regards classification as one of the primary prerequisites in the advancement of knowledge even though it recognizes that there are problems and dangers in the misapplication

¹ Chief, Laboratory of Clinical Science, National Institute of Mental Health, Bethesda, Md.

of classification. I should like to examine the needs for and the processes of classification in biology, and their implications for psychiatry, even though much of what I shall say may be self-evident by now.

Fundamentally, classification is an essential part of the process by which science comes to grips with the universe. It is essential in structuring the search for more information and in meaningfully and effectively communicating what has been learned; for these reasons it must occur very early in the scientific process.

It is quite clearly analogous to the way in which an individual learns to comprehend the universe. First, there are phenomena impinging on all the senses in what must be a hopeless and chaotic array. Yet, gradually and by experience and testing, these phenomena are seen to cluster, to associate themselves into groupings, which, when tested by the child himself and confirmed by the independent experience of other individuals, testify to the existence of an outside world and of specific entities such as milk, apples, or books.

I should like to examine two examples of classification in biology and discuss how they have influenced the field. The first is the old and wornout example of Linnaeus and his classification first of plants, then of all of the kingdoms of nature.

Linnaeus, although he was a botanist and made his important contributions to botany, was also a physician. As a matter of fact, among the other things which he attempted to classify were human diseases. Indeed, it might be interesting to look at his "Genera Morborum" in view of his success in other areas.

Early in his life he became impressed with the significance of stamens and pistils (thus reinforcing an important psychiatric insight), and developed the idea of classifying plants by means of those structures. It is rather interesting that Linnaeus had a strong belief in the fixity of species, which turned out eventually to be wrong, which motivated his attempts at classification but

did not disturb his observations. Nor did the Linnaean system prevent Darwin, who began as a taxonomist and classifier, from developing his very important, imaginative concepts; in fact, he could not have done so without it. It is not a controversial assertion to say that the Linnaean classification made possible not only the imaginative theory of Darwin but was also responsible for much of the future development of biology.

There is another area in biology, more closely related to our problem, perhaps, in which classification has been of importance—the field of pharmacology.

The Chinese pharmacopeia, which I suppose is rapidly being supplanted by occidental systems, was developed by the emperor who tasted all of the existing medicinal herbs and classified them into weak, strong, and intermediate.

In the Western World, classification proceeded somewhat more slowly. The present system which has been developed over hundreds of years, is, interestingly enough, also a phenomenological classification. Restricting ourselves to the drugs which act upon behavior, we have anesthetics, hypnotics, sedatives, analgesics, analeptics, and convulsants. This has not precluded the addition of new classes, such as ataractics and antidepressants, or further subdivisions of these phenomenological classifications according to further evidence (which is nevertheless still phenomenological) after such evidence has been rigorously established. Thus, we have subdivisions in terms of chemical structure, site or mechanism of action, specific therapeutic use, to name a few.

It is interesting that in pharmacology as in most systems of classification and nomenclature, these categories are not always mutually exclusive. Chlorpromazine, for example, may have to be included under both ataractics and antiemetic drugs. Nor can any system always be correct; as a matter of fact, digitalis was first classed simply as a diuretic drug, before it was realized that its primary effects were upon the heart.

With these two examples in mind, it may be possible to set down certain requirements of a classification system useful for research in biology. I suspect that these same general principles may be applicable to psychiatry, for even though one's interest may be in facilitating research and the acquisition and communication of knowledge, it is likely that a system which fulfills these aims will, with remarkably little change, fulfill many of the other purposes which we may identify in psychiatric nosology.

In the first place, systems of classification in biology begin with an implicit assumption that there is order and differentiation in the universe under examination; I would submit that from the experience of biology and of other fields, this seems a worthwhile and heuristic assumption to make in the field of human behavior.

It may develop that there are large segments of human behavior, large areas of what are now regarded as mental diseases, which will never permit differentiation, but somehow I feel that this will be demonstrated best by testing the validity or demonstrating the spuriousness of the obvious phenomenological clusters rather than rejecting them *a priori* on a hypothetical or wish-fulfilling basis.

Secondly, although these classification systems in biology have been frankly and blatantly phenomenological, they have somehow insisted upon maintaining standards of what constitutes valid phenomena and primary data and what differentiates these from inferences and hypotheses.

For example, before infectious processes were admitted to etiology and became accepted in the classification of medical disease, there were Koch's postulates to be fulfilled, which were fairly rigorous requirements of what constitutes etiological evidence.

In certain areas of biological psychiatry, there are subdivisions which were developed on the basis of new phenomenological observations such as phenylketonuria. Here again, before these became useful categories

for classification, certain standards were met regarding the existence of such a subgroup, the common features of its members and the specificity of the new phenomenon, i.e. the excretion of phenylpyruvic acid and other phenyl ketones. It is important that we do not too easily form subgroups in psychiatry each time a new component is reported as occurring in the urine of certain types of mental patients.

The third requirement of a useful classification system, I feel, is that it does not permit inferences which beg important questions and which it is the very purpose of the classification to help solve. In other words, if one includes in a definition of schizophrenia inferences of its genetic or experiential nature, then one has begged the question and one will inevitably find evidence supporting one or the other factor in schizophrenia.

If one includes in the definition of certain types of psychosis an inference that they are psychogenic, as, for example, the category psychogenic psychosis which exists in certain systems of nomenclature, one has already answered an etiological question which a system of classification should help to pose.

The problem between endogenous and reactive depression, I suggest, is a problem which comes from a failure to observe this important rule. Until definitive evidence for the endogenous or reactive nature of a condition is at hand, it may be well to restrict the nomenclature entirely to the phenomenology of the state rather than to inferences regarding its etiology. An effective nosological system should stimulate inferences and aid in their testing but not incorporate them until they are established by accepted rules of scientific evidence.

A fourth characteristic is that any system of classification or nomenclature must recognize and be able to entertain uncertainty. It must be flexible, fluid, and tentative. It must be a continuous process dependent upon the phenomenology at hand as elicited by current procedures and techniques. It

must therefore be capable of revising old categories, recognizing new classes and subclasses in the light of new evidence.

I think such a requirement could avoid the premature reification of knowledge by the very process of classification, could prevent dogmatic inferences which arise from the system of classification *per se*. If psychiatry recognizes, at the fundamental level of nosology, that the great problems with which it wrestles, such as schizophrenia or depression, are simply phenomenological clusters with little demonstrated etiological or pathogenetic meaning, it may keep itself from a premature adoption of whatever etiological hypotheses happen to be in vogue, whether they be organic, psychodynamic, or psychosocial. Nosology may then serve best its important purposes of defining and communicating what is known, indicating what is unclear and stimulating the acquisition of evidence in those areas.

There are many problems in psychiatry which are much more formidable than those which the rest of biology has ever faced and one can enumerate certain of them:

There is a considerably larger subjective element in the primary data of psychiatry than in that of the other branches of medi-

cine. Radiologic evidence for pulmonary fibrosis is considerably more objective than the indications of affective state even though the interpretation of both types of evidence requires subjective judgments. Psychiatry also deals more with functional rather than structural alterations in underlying processes, and functional changes are more difficult to observe and to validate than are structural alterations. Interpersonal relationships are of considerably greater importance in psychiatry than they are in other branches of medicine, and these are not easy to define or control.

There is what appears to be the important distinction between the digital data of biology and the analog data of psychiatry. I can't help but feel that these problems, which are real and which undoubtedly help to explain the somewhat slower pace at which classification has occurred in psychiatry than in other branches of medicine or biology, do not constitute qualitatively different problems from those which are faced in these other areas, and that somehow the basic principles of nosology, nomenclature, and classification which have been found useful in other sciences are equally useful in psychiatry.

ISSUES IN THE METHODOLOGY AND STATISTICS OF CLASSIFICATION

The several steps in the development of a classification system are reviewed and clarified by mathematical statisticians and psychologists. The major issues with regard to the selection of appropriate statistical methods for each of these steps are identified. Progress in the refinement of methods for dealing with the problems of discrimination, measurement of similarity between profiles, mixture, classification, and clustering are reviewed. Several focal issues were given special attention in the discussion. These included consideration of the appropriate methods for determining the underlying structure of multivariate data, and the distinction between the roles of descriptive and probabilistic techniques in identifying patient clusters.

On the Meaning of Discrimination, Classification, Mixture, and Clustering in Statistics

Samuel W. Greenhouse, Ph. D.¹

Introduction

My purpose is to define the various terms, and hence the problems, occurring in statistical theory and methodology in connection with the subject matter of this conference. The reason for doing so is to bring before the participants what statisticians or psychometricians tend to mean by discrimination, classification, and clustering. These meanings perhaps do not precisely coincide with those definitions attributed to these terms by others. Furthermore, it is not clear whether even all statisticians would agree with the definitions put forth below. Defining terms in any field to the satisfaction of all is a very difficult enterprise. This is particularly true in the behavioral sciences so that I am certain the psychiatrists and clinical psychologists present will sympathize with me.

Discrimination

In statistics, the problem of discrimination is in a certain sense prior to that of classification. In most contexts, statistical discrimination refers to the following question: Do these two populations differ? In connection with the classification problem, statistical discrimination refers to the question: Do these two populations differ to

the degree where classification is worthwhile? Here populations, of course, relate to frequency or probability distributions of random variables. This is not as unrelated to real world populations as it may seem.

For example, consider the trait "mental status" of an individual and assume the existence of some measurement instrument designed to measure this trait. The application of this measuring device to the i^{th} individual yields a numerical score x_i . Clearly, different individuals will yield different scores. Thus, we may postulate the existence of a frequency distribution of scores in a population of individuals, $f(x)$, say. We denote $f(x)$ as the frequency function or probability density function. We call $F(x)$ a probability distribution in this sense: for any value of x , x_0 , say, we can compute the probability that an individual will have a score less than or equal to x_0 and that there exists some value of x such that all individuals in the population have values less than it. It is common for the research worker or the statistician to specify a reasonable mathematical form for f . In any given population, the probability function, $f(x)$, possesses certain properties, usually referred to as parameters, such as measures of location (means) and of scale (variance or standard deviation). If θ represents the vector of such parameters, then we denote the distribution of x in the population by $f(x; \theta)$.

¹ Chief, Theoretical Statistics and Mathematics Section, Office of Biometry, National Institute of Mental Health, Bethesda, Md.

The discrimination problem then arises in the following way. There are two or more groups of individuals (humans, flowers, mice, etc.) which are known to differ by some readily recognizable characteristic other than x . The questions that are of interest are the following:

(1) Is x distributed in the different populations (assume two, for simplicity) according to the same distribution form? That is, is the form of $f(x)$ the same in both populations?

(2) A more common situation is to assume that the functional forms are the same and ask whether the parameters are the same; i.e., is $\theta_1 = \theta_2$?

(3) Finally, for purposes of classification, it is important to ask whether x separates the two populations far enough apart so that classifying individuals on the basis of x is worthwhile.

I believe it is now clear how the problem of discrimination is prior to that of classification. If, on the basis of a measurement x , we are to classify or assign an individual to one of two populations, we should like to know that those two populations are indeed different. In a more general sense, just this question by itself has given rise to that body of statistical theory known classically as "hypothesis testing." However, as a part of the classification problem we would like to know more than simply whether $f(x; \theta_1)$ differs from $f(x; \theta_2)$. If they differ by a "little" clearly the classification that follows may be very poor in that too many errors or misclassifications occur. Thus, it is desirable to know how well does x discriminate between the two groups. Further, it should be added that in a vast majority of such problems x is not one variable but represents a vector of many measurements. There was then a need to develop measures to reflect the difference between populations which are known as distance functions. It is not generally appreciated that the square of the t statistic used in testing the difference between two samples means is a constant times a distance function between two samples.

Similarly, in the case of multiple measurements, investigations centered on reducing the problem from multidimensional space to a single variate by finding optimum functions of the variables. This constitutes the study of discriminant functions. Research in these areas is by no means complete and many questions still remain to be resolved.

Classification

In some references, the term "assignment" is used. I believe I am correct in stating that invariably the assignment problem is identical with the classification problem as defined below.

The problem of classification, as it appears in statistics, can be described as follows. There are two populations each defined by $f(x; \theta_k)$. It is assumed the degree of discrimination between the populations on the basis of x is satisfactory. An individual is assumed to belong to one or the other of the two groups but which one is unknown. An observation is made yielding the vector x . How can one best assign him to one of the two populations?

There exist several different solutions to this problem in the statistical literature. The classical solutions based on discriminant functions have been based on the following rule: if the function of the measurements is denoted by $y(x)$, so that the value of the function for the observational vector, $x_{(i)}$, is y_i , assign the individual to one population if $y_i > y_0$ and to the other population if $y_i < y_0$ where the functions $y(x)$ and the "critical" value y_0 are determined in the discrimination solution as giving the greatest discrimination between $f(x; \theta_1)$ and $f(x; \theta_2)$. Most of the work here has been with probability functions which are assumed to be multivariate normal and with discriminant functions which are linear combinations of the measurements. Another solution to classifying the individual makes use of the likelihood function of the observation and in effect computes a likelihood ratio odds. This is simply $f(x_{(i)}; \theta_1)/f(x_{(i)}; \theta_2)$. If

the ratio is greater than some number c , one would assign the individual to the first population; if the ratio is less than c , then one would assign him to the second population. (As presented here, these constitute two-decision rules. It is clear that these can be extended to three decision rules by introducing a range of values, such as $\frac{1}{c} < f(x_{(i)}; \theta_1)/f(x_{(i)}; \theta_2) < c$, for example, in which no assignment would be made).

The two procedures presented are not unrelated. In fact, optimum discrimination functions are yielded by likelihood ratios.

More recently, with the increased interest in Bayesian inference, some work has been done in finding the probability in the posterior distribution that an individual belongs to one of the populations. The investigator will then classify by comparing the various posterior probabilities available for the individual. Here one must assume a prior probability function for the unknown parameters θ .

Finally, it should be pointed out that many statisticians do not make the separation between the two problems of discrimination and classification implied here. Often they consider one problem under the heading of classification. I note with interest and some satisfaction that Professor Rao, according to the title of his paper, does separate these problems and will speak to us shortly on them.

With reference to the lack of unanimity on the definitions made here, at a conference on multivariate analysis held in the spring of this year, M. G. Kendall, if I heard him correctly, denoted the operation of separating populations and determining by how much they differ as the classification problem and he referred to the process of assigning an individual observation to a population as discrimination.

Mixture

This terminology is probably the least known to those outside the statistical profession. The mixture problem considers a

sample, usually large, as composed of individuals from several different populations. It is not known which individuals come from which populations. The most general form arises when the solution requires estimating both the number of populations concerned and the identification, by estimation of the various θ involved, of the different populations. A special case is when one can assume the number of groups represented in the sample and must then identify and estimate the different groups.

This problem is very difficult and very little work of a probabilistic nature has been done in the past. Most investigators confronted with this question, particularly in the behavioral sciences, have partitioned their sample by eye, so to speak, using descriptive principles which were not based on any probabilistic grounds. It will be noted that partitioning the sample into clusters or "constellations" as they are called by Rao is also part of the clustering problem to be discussed below. Also one will note the similarity of the problem to that of factor analysis where one desires to reduce a set of variables into a smaller number of psychologically meaningful constructs where the number and identification of those factors are unknown.

The availability of high speed electronic computers has made possible the beginnings of an effective, probabilistic attack on this problem. Later in this session, we shall be privileged to hear Professor Barnard describe the work he has done in this area.

Clustering

The problem of clustering is not completely distinct from the problems previously discussed. It is a term which originated with psychologists and many of the procedures used in behavioral science research have been developed by them.

To avoid confusion I must point out a problem appearing in statistics known as clustering which is somewhat different from what we are discussing here. This has to

do with the detection of spatial and temporal clusters of events when such exist, and arises frequently in the epidemiology of diseases.

Clustering as used in the behavioral sciences has two major subdivisions. One pertains to the clustering of individuals or more precisely of sampling units into more or less homogeneous groupings. The other pertains to the clustering of variables, traits or measurements.

The former problem is related to the mixture problem in the same sense as classification is related to discrimination. If a reasonably good solution were to exist in the mixture problem so that we could conclude it was most probable that a given number of populations were represented and these could be reasonably identified through the estimation of the parameters, one could then cluster individuals by classifying them into one of the determined populations in accordance with some optimal probabilistic procedure.

The development of procedures for clustering individuals, however, did not and could not wait for mathematical statisticians to solve the mixture problem. As a result, there are many descriptive, empirical methods that have been proposed. The criteria behind some of these range from very simple, obvious principles to very complicated and sophisticated ones. Some of the obvious criteria for clustering people appear in very early literature, such as the Old Testament, where the delineation is made on the basis of carefully described symptoms of certain diseases. In the beginnings of modern statistical practice in the latter part of the last century, Karl Pearson proposed a criterion which has become known as the coefficient of racial likeness. More recently we have had the Cronbach-Gleser distance function which was designed to cluster individuals by their similarity two at a time. Again, we shall be fortunate to hear Professor Gleser discuss this question of measuring the similarity between individuals during this session.

The other subdivision in clustering is that of finding clusters of measurements. This problem is quite distinct from that of clustering individuals. In statistical theory, much attention has been devoted to this problem in multi-variate analysis. In my own view, it is important to distinguish between two different objectives of the investigator. In one case, the purpose is simply to reduce a very large number of observations to a smaller number with a minimal loss of information. The principle of parsimony here is not so much prompted by theoretical considerations as it is by operational and computational simplicity. The other aim demands much more of the clustering technique. It seeks to cluster variables into homogeneous sets each of which reflects some underlying trait or capacity, having some psychological significance, which is hypothesized as existing but which is not directly measurable itself.

There are quite a few techniques available to accomplish these objectives. Statisticians have developed principal components, canonical correlations and other procedures based on taking linear combinations of the original measurements. Psychologists and psychometricians have developed factor analysis, in all its various forms, and in addition many empirical clustering techniques based on distance like functions or on correlation coefficients.

I assume Dr. Lyster will discuss many of the problems inherent in techniques for clustering both people and variables.

Conclusion

Despite the necessity of having to enter into discussions of detail in order to define the various terms with which we are concerned, it seems to me that we must not lose sight of the overall objective of this conference, namely, "to explore the roles currently served by diagnosis and other classification procedures in various aspects of clinical psychiatry and research in psychopathology." I have heard it said on innu-

merable occasions that the diagnosis put on an individual entering a mental institution is of very little consequence; in fact, that this is left to the nurse or some other attendant to do, and that the therapy and course of action to be followed are determined by things other than diagnosis. To the extent this is true, it merely reflects ignorance of the disease process or simply a bad nosology. Everyone seems to be agreed that the clinician will treat symptoms. But what is apparently happening in psychiatry is that there does not exist a fair degree of agreement on the part of clinicians as to the correspondence of symptoms with disease categories. Maybe this is inherent in the nature of behavioral and social disorders. Perhaps, also, it is unnecessary to clarify this situation in the treatment of individual cases. But it seems to me this situation is catastrophic in the progress of research in the mental disorders. If the scientist cannot be sure of the classification procedures which lead to the groups of individuals he is studying, it is not clear what further advances can be expected.

This leads me to comment that no conference on classification can be complete without some reference to Linnaeus, even if the conference does not deal with the plant kingdom. In his "Genera Plantarum," first published in 1737, we find the following remarks:

"All the real knowledge, which we possess, depends on METHOD; by which we distinguish the similar from the dissimilar. The greater number of natural distinctions this method comprehends, the clearer becomes our idea of the things. The more numerous the objects, which employ our attention, the more difficult it becomes to form such a method; and the more necessary."

"For we must not join in the same genus the horse and the swine, tho' both species had been one-hoof'd, nor separate in different genera the goat, the reindeer, and the elk, tho' they differ in the form of their horns. We ought therefore by attentive and dili-

gent observation to determine the limits of the genera, since they cannot be determin'd a priori. This is the great work, the important labour, 'for should the Genera be confused, all would be confusion.'"

Quantifying Similarity Between People

*Goldine C. Gleser, Ph. D.*¹

Clinical and counseling psychologists are often concerned with the problem of comparing persons on the basis of a number of test scores or ratings considered simultaneously. One of the purposes for such comparisons is the identification of a limited number of "types" of people having similar configurations of scores. In order to make such groupings it is necessary for the investigator to choose an index of similarity from a large number of possible measures. Each measure defines similarity somewhat differently, and hence leads to the identification of different configurations of scores as maximally similar. The empiricists, not discouraged by the multiplicity of indices, have often tried different indices on the same data in the hope that the appropriate measure of similarity would reveal a natural clustering of persons that could form the basis for a new and more meaningful typology.

Perhaps it is time for investigators to discontinue trial-and-error methods and to reconsider their reasons for developing typologies in the light of known principles of measurement and the assessment of individual differences. It is my contention that typologies are unnecessary where more refined measures are available for decision making. The determination of types by means of similarity indices and clustering

¹ Professor of Psychology, Department of Psychiatry, University of Cincinnati College of Medicine, Ohio.

procedures is simply an arbitrary repartitioning of the measurement domain into coarser groupings with a consequent loss of information. Such grouping in most cases will lead to poorer decisions than would be possible if the same measurement data were utilized in more efficient ways.

Some General Considerations Regarding Typologies

Attempts to devise typologies of people has had a long history. According to Rees (13), the ancient Hindus used sex, physical, and behavioral characteristics to classify people into six types which they designated by names of animals. The early Greek and Roman physicians developed several typologies based on variations in physical characteristics presumed to stem from a mixture of the four humours. The most prominent of these typologies was that of Galen (ca. A.D. 129–199). He defined nine temperamental types which were assumed to relate to a person's susceptibility to various diseases and to individual differences in behavior.

Essentially types can be thought of as a set of mutually exclusive, qualitatively different categories constituting a nominal scale. They are a way of abstracting and condensing information about the myriad behaviors of people into a form which can be handled more easily, grouping together into one category those persons who are minimally dissimilar in some set of characteristics deemed to be important to the understanding and prediction of behavior, and identifying in some fashion a limited number of categories to account for the totality of individual differences in the domain under consideration.

Where nominal scales have been formulated to differentiate among persons (or stimuli), it is often found later that these and finer differentiations can be made by determining some one or more dimensions which account for the grouping. Thus colors which appear qualitatively different

can be described quantitatively according to their component wave lengths and the intensity of light at each wave length. Similarly, the body types of Kretschmer and Sheldon are now being replaced by continuous indices of body build and body size (13).

Similarity as a construct.—"Similarity" is a construct which is at the basis of all measurement—particularly the measurement of individual differences. When we place objects in one-to-one correspondence with a set of scale points we are implying that objects assigned to the same point are alike in some respect while those assigned to adjacent points are similar. For example, a rating scale is a method of matching persons to a set of descriptions so that persons considered most similar in the pertinent attitude or behavior are grouped in the same category. If the scale is ordinal, persons in adjacent categories are more similar than persons in categories further removed. In an interval scale the arithmetic difference between numbers assigned to categories represents the extent to which persons are viewed as dissimilar. It is important to note that the extent of similarity or dissimilarity is always indicated with respect to some characteristic, trait, or dimension. Persons falling in even the two most extreme categories of some scales can be seen as similar or alike in terms of another construct. For example, on the scale "time spent sleeping" the scale points "almost always sleeping" and "can't ever sleep" are maximally dissimilar. However, from the standpoint of degree of illness they could be considered identical.

When one is dealing with a complex description such as is embodied in a profile or set of scores, it is particularly necessary to specify what is meant by similarity. Are differences between persons along each dimension of the profile counted or is it only the patterning of score elements that is of importance? Each definition will lead to a different measure of similarity and consequently to a different grouping of persons.

The generality of types.—When the question is asked, “How should profile similarity be measured?” it seems to imply that there is some “true” or “right” way to classify people inherent in the data. This notion is an oversimplification which violates the fundamental principle that a “type” is simply a cognitive construct by which we condense information. There is no one meaningful way to classify people; nor will people alike with respect to one set of attributes or scales necessarily be more similar to each other than to persons in general on other characteristics. People can be classified into types in as many ways as we can specify the behavioral basis or the physical, mental, or emotional characteristics for making the judgment of similarity. Each such classification will probably relate to some aspects of the person’s past, present, or future behavior and not to other aspects. For example, intelligence is a concept, embodying a complex variety of behaviors, which can be used to classify persons. Similarity in level of intelligence is related to similarity in interests, educational achievement, occupational choice, economic status, etc. However, persons of similar intelligence may vary as much in character, personality, or response to drugs as do people in general. Different scales or classifications are needed to predict these aspects of the person. Any particular classification will be meaningful only to the extent that it is based on variables or attributes which are related to the broader class of behavior one is trying to predict or control. In psychiatry it is very likely that many classifications will be needed. If these are typologies, the problem again presents itself of multiple measures on each person, but of a categorical, unordered kind, less amenable to powerful methods of statistical analysis.

Loss of information.—Another fallacy often implied is that a method of grouping persons by similarity measures can be found that does not discard important in-

formation. However, the question of whether or not information is important depends on whether it is pertinent to the behavior we want to predict. Actually, any encoding of behavioral phenomena discards some information that might be important. There is no way to structure phenomena into characteristics, variables, or typologies without some condensation and generalization. It is evident that summing the number of correct items or the number of symptoms of a particular kind to form a scale discards information. If one has simply noted the presence or absence of 10 characteristics, there are potentially 2^{10} or 1024 different patterns of such characteristics, so that 1024 differential predictions could be made in place of the 11 differentiations made by a simple frequency scale. The frequency scale implies that persons having the same number of symptoms of a certain kind are considered similar for decision purposes. Information regarding the patterning of symptoms is discarded. In a like manner, if a different definition of similarity is used to combine patterns of symptoms, the resulting classification would most likely contain a limited number of subgroups with some corresponding loss of information. The important question to consider is what information is relevant and what is expendable. In order to answer this question the investigator must have some knowledge of what he means by a good classification—either in the form of clinical judgment or of external criteria.

If, instead of attributes or other binary data, 5-point rating scales are used to describe behavior, then, with just 5 variables, 3,125 score configurations are possible. Reducing these potential discriminations to, say, 10 or 12 types, results in a considerable loss of information unless the probability of occurrence of most of these configurations is extremely low. Actually the difficulty here is that, despite the condensation accomplished by the scaling, the investigator has more information than he

can easily comprehend or handle. Hence, the easier route is taken of neatly pigeon-holing people, rather than dealing directly with the multivariate information.

A Geometric Model of Profile Similarity

Let us consider a geometric model for the concept of similarity between persons that facilitates comparison among the various indices commonly used. Assuming that each person has been observed on k variables, he has a set of scores which can be designated

$$\{X_{p1}, X_{p2}, \dots, X_{pk}\}.$$

If we regard each variable as defining one of the k mutually orthogonal axes, this set of scores can be represented as a point in k -dimensional space. Points will be scattered with varying concentration throughout this space, their configuration determining the marginal distribution of scores on each variable and the intercorrelations among the scores. It should be evident that for two variables this representation is identical to the usual scatter diagram as illustrated in figures 1 and 2. In some cases the configuration of points is contained in a subspace which can be described by less than k dimensions. This situation occurs whenever there are constraints on

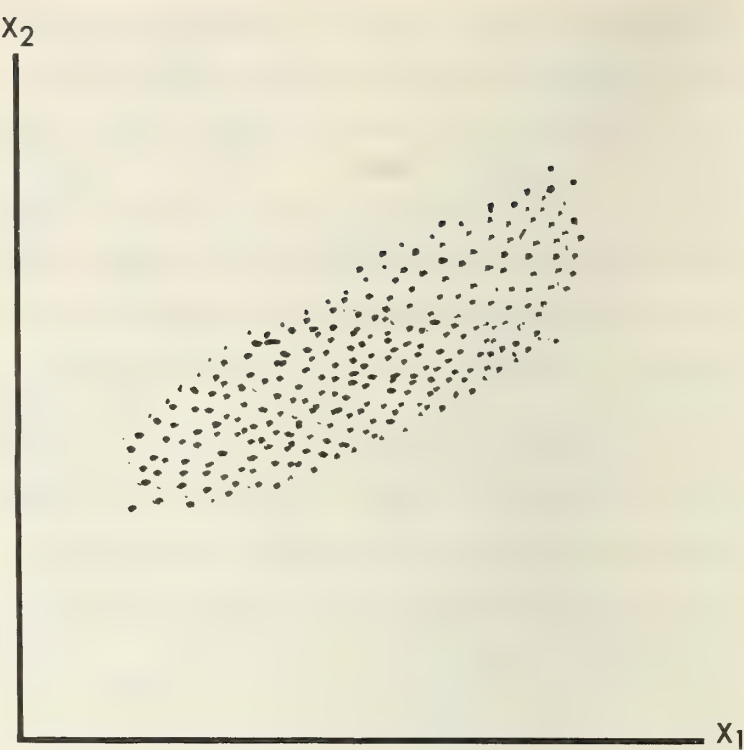


FIGURE 1.—Bivariate normal distribution.

the sets of elements such that some elements can be determined from linear combinations of other elements. We shall consider this possibility in more detail later.

Consider, now, the two-dimensional scatter-diagrams illustrated in figures 1 and 2. The first figure is intended to represent the bivariate normal distribution usually assumed in statistical inference. If such a distribution were actually approximated by the data the sample would be considered homogeneous and no subgrouping would be called for. The concept of types which

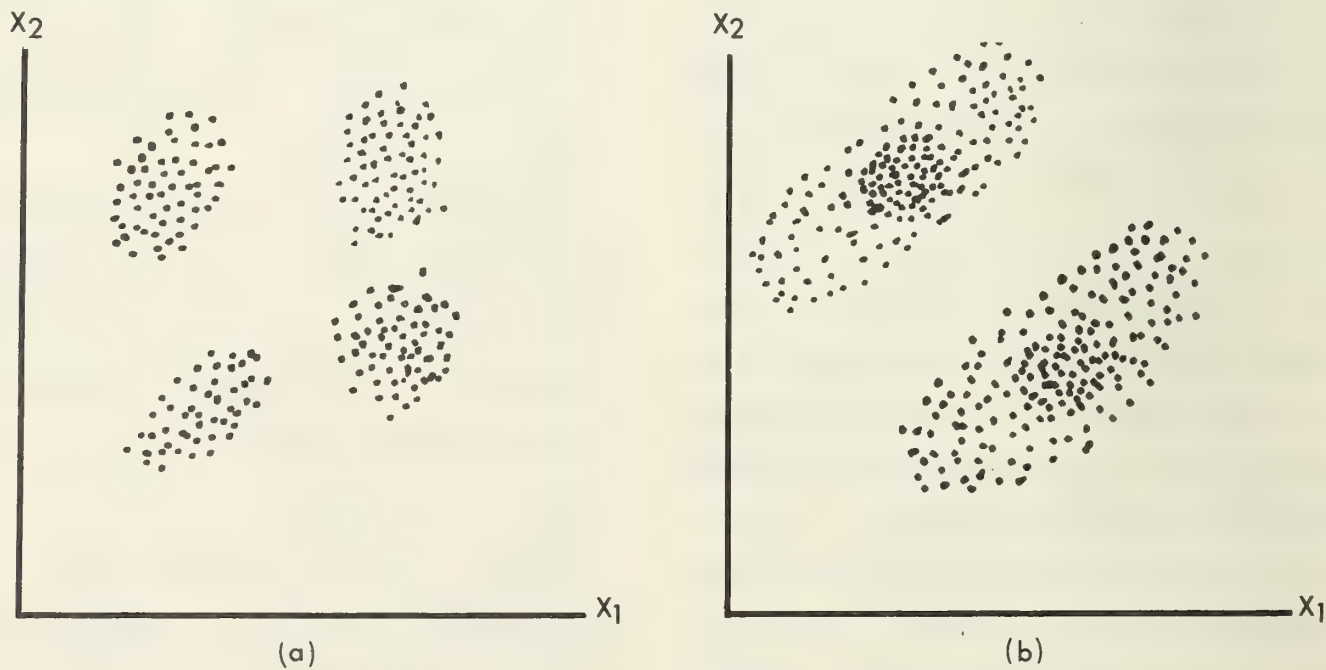


FIGURE 2.—Clustered distributions.

most clinicians seem to hold is more nearly represented by 2a or 2b where points lie in nodular clusters. Figure 2b, incidentally, represents one of the favorite typologies of clinicians; a subgroup for which $X_2 > X_1$ and another for which $X_2 < X_1$. Interestingly enough, approximately normal marginal distributions are possible for constellations such as those shown in both 2a and 2b.

Unfortunately, since concern is usually with more than two dimensions, visual inspection of the data becomes well-nigh impossible. Therefore some measure of similarity is chosen which hopefully will reveal meaningful clusters when those profiles that are most similar according to the index are grouped together. Many of the indices of profile similarity treat all the points in certain configurations of the total space as identical and hence may be thought of as transforming all points lying in this configuration to a single point. Adjacent configurations in the original space become adjacent points in a transformed space. The new set of points can then be described by fewer dimensions than were needed for the original complete set. Thus, such indices, by ignoring some source of dissimilarity in the original configuration automatically cluster or group persons to some extent. The choice of such an index reflects assumptions concerning the lack of importance of certain types of differences.

Characteristics of Specific Similarity Indices

We shall consider some of the more frequently used measures of profile similarity, illustrating them whenever possible by two-dimensional figures.

The D measure.—The distance measure suggested by Cronbach and Gleser (2) and independently by Osgood and Suci (10) defines dissimilarity between two profiles as the Euclidean distance between the two points representing these profiles in

k -dimensional space. Two profiles are considered identical if and only if the points are coincident, that is, if the two sets of scores are identical. The greater the distance between two points the more dissimilar are the two profiles. Either D^2 or its square root, D , can be used as the index; D^2 being determined for the points X_{pj} and X_{qj} as

$$D_{pq}^2 = \sum_{j=1}^k (X_{pj} - X_{qj})^2 \quad (1)$$

The square root of this measure, D , has considerable logical appeal since it seems to preserve all the information in the original measurements intact without making additional assumptions. Assuming that a nucleus were chosen in some manner the D index would treat all points lying in a concentric shell around the nucleus as being equally similar to the nucleus. While either D or D^2 might be used as an index of similarity, D appears more appropriate for clustering procedures.

The distance between any two points in space is determined by the degree of discrimination along each variable. A change of units in any direction will modify the distance. Quite often the fineness of quantification of any particular variate depends on the extent to which we can make discriminations, or in the case of test data, on the items we happen to use. However, the problem here is not simply a matter of the arbitrary nature of the psychological units, a difficulty which can often be solved by standardizing variables as suggested by Overall (11). The difficulty is that discriminations, on any one scale or on different scales, may not necessarily be of equal importance to our concept of similarity or to our classification purpose. For example, an elevated temperature enables one to diagnose the presence of an infection of some kind, but the only additional temperature discrimination usually helpful in medical diagnosis is the differentiation of high-grade and low-grade fever. If finer

discriminations are made on a variable than is necessary, and persons are grouped by presently available similarity indices, each such discrimination will automatically be treated as equally important to every other discrimination, whether on the same or different scales. Thus, one degree of difference in temperature would be treated as equal to one unit of intensity of headache, etc. The taxonomist usually has some weighting in mind in making hierarchical classifications, but it must be specified if it is to be reproduced in the statistical model. As Rogers and Tanimoto (14) have emphasized, clustering techniques can lead to meaningful classifications only if the scientist can decide on the relative importance of different attributes to his total problem. A variation of D which might be useful in such situations is given by

$$D_{pq}^2 = \sum_{j=1}^k w_j (X_{pj} - X_{qj})^2 \quad (2)$$

where the w_j are differential weights applied to the variables to represent a value-judgment regarding the relative importance of each variable in the overall discrimination. This would have the effect of stretching (or shrinking) certain dimensions of the original space relative to the others.

It should be fairly evident that if Euclidean distance is used as a measure of similarity, then the average squared distance between pairs of points in a group and the average squared distance of such points from their common centroid (D_{op}^2) are both measures of the heterogeneity or dispersion of the profiles. Thus

$$D_{pq}^2 = 2 \frac{N-1}{N} \sum_j s_j^2 \quad (3)$$

and

$$D_{op}^2 = \frac{N-1}{N} \sum_j s_j^2 \quad (4)$$

where s_j^2 is the unbiased variance of element j in the sample or subsample.

Ward (16) looks upon the dispersion or dissimilarity given in equation 4 as the loss of information when the average profile of a group of persons is used in place of their individual score sets. Treating each score set as unique yields maximum information whereas clustering reduces information to the extent that D_{op}^2 over all clusters differs from zero. Profiles can then be clustered to yield the minimum loss of information at each stage. This approach emphasizes the point made earlier that typing results in loss of information. One must decide on other grounds what kinds of information are expendable and the amount of loss that can be tolerated.

One of the questions that has often arisen in connection with using D as the definition of profile similarity is whether the measure is appropriate or meaningful if the variables constituting the profiles are correlated. Overall (11), for example, claims that the Pythagorean distance model requires comparable scale units and zero correlations among profile elements and hence recommends using the distance measured in a space such that the principle components of the original variables are the orthogonal axes and scores are standardized on each factor. In terms of figure 1, this space consists of a rotation of axes from the ones illustrated to a new set roughly corresponding to the major and minor axes of the ellipse formed by the cluster of points, and then a stretching or weighting of the minor axis with respect to the major axis so that the cluster of points becomes circular.

It can be shown that distances between profiles of standardized test scores yield the same value as would be obtained by using equation (2) on principal component factor scores, with w_j equal to the latent roots or eigenvalues. The D^2 measure weights the principal component factor scores according to the extent that they are represented in the original set of variables. Whether this or some other weighting is preferable is a matter for the investigator

to decide on a priori or empirical grounds. Neither distance, as Heerman (5) points out, can be considered a true distance nor is either distorted.

The choice between factor scores and variable scores becomes even more perplexing when a distance measure is used to identify clusters or types. The assumption that clusters are present implies that the sample at hand has been obtained from a heterogeneous population consisting of several distinct populations or subgroups. If this is so, the correlations among the scores in the total sample reflect to some extent the configuration of subgroup means. The appropriate factor space for the Mahalanobis (7) distance measure recommended by Overall is that of the within-group variance-covariance matrix. But this cannot be determined unless the subgroups can be identified by some independent consideration. One possible solution is to assume that within the subgroups the correlations will be negligible, an assumption similar to that of Lazarsfeld's (6) structural analysis. Another possibility would be to determine the factors from scores of a sample drawn from some known homogeneous population, and use this information to convert raw scores into factor scores in the sample to be investigated.

While it is not necessary to have uncorrelated variables in order to consider D^2 as an index of similarity, it is probably desirable that the correlations be low. If correlations are high, the first few principle component factors account for almost all of the dispersion and hence there would be little or no loss of information if the profile elements were replaced by a fewer number of factor scores. Geometrically, this means that the original configuration of points can be contained almost completely in a space of considerably less than k dimensions. The principal component factor scores, consisting of linear combinations of the original scores, effect the transformation to this smaller space, collapsing certain closely spaced points on

dimensions orthogonal to the factors into a single point. If profiles are reduced in this manner, it is still necessary to decide what weighting to give each factor in any further treatment of the scores.

The index r_p .—An interesting index of profile similarity which is a function of D^2 was proposed by Cattell in 1949 (1) and is currently advocated by him for comparing 16 PF profiles. This index may be written as

$$r_p = \frac{2_k \chi^2_{50} \sum s_j^2 - k D^2}{2_k \chi^2_{50} \sum s_j^2 + k D^2}$$

where ${}_k \chi^2_{50}$ represents the median χ^2 corresponding to the given number of variates. The index ranges from $+1$ to -1 as D^2 varies from 0 to ∞ , and its expected value is 0 for large values of k and N . r_p is a nonlinear monotonically decreasing function of D^2 which, relative to that measure, tends to discriminate sharply between rather similar profiles, but treats all deviant profiles as of roughly equal dissimilarity. So far as I know this index has not been used for classifying profiles. Nunnally (9) points out that the matrix of r_p 's between all persons in a sample is non-gramian and hence not amenable to the usual techniques of factor analysis. However, there is no reason why it could not be used in conjunction with clustering procedures.

Measures which reduce the dimensionality of dissimilarity.—Let us now consider some indices which equate certain subsets of points in space and hence reduce the dimensionality of dissimilarity. Such measures, by their nature, perform a clustering or grouping function that should be recognized by the investigator when he chooses an index by which to compare or cluster profiles.

The most common index of this type is the correlation between profiles, or Q correlation (15), although in some cases the covariance between profiles has been employed (8). It is a well-known fact that both these measures remove any variance

due to elevation, and hence reduce the dimensionality of the space to $k-1$. "Elevation," which is the average of the elements in a profile, very nearly corresponds to the first principal component of the variables. This factor is ignored as a source of dissimilarity when either covariance or correlation between profiles is used as the index of similarity. In addition, the correlation measure removes the variance due to "scatter" in the profile, thus projecting the point representation of the profile to a $k-2$ space.

"Scatter" here refers to the standard deviation of score elements which varies from individual to individual. These matters are probably well known to many of you, but let me try to illustrate what it means for those who are less familiar with the notion. In figure 3 several points are plotted representing profiles of two elements ($k=2$). Removal of elevation, by subtracting the person's mean score from each raw score, projects these points on to the line L_1L_2 . Thus the points are now constrained to a one-dimensional space. All points lying on a line perpendicular to L_1L_2 are collapsed into a single point: The profiles they represent are treated as identical. Thus dissimilarity along this dimension is ignored.

In order to illustrate scatter, let us assume we started with a three-variable pro-

file and then proceeded to remove elevation. The profiles can now be represented by points in all four quadrants of a two-dimensional space, the centroid having been translated to the origin. The scatter or standard deviation of scores around the mean of any particular profile can be represented by the distance from the origin to the point representing the profile. If scores are standardized within profiles, all points are projected to a circle at a unit distance from the origin. Points lying anywhere along a particular radius will be treated as identical by any measure which equates scatter. (See figure 4.) Dissimilarity is measured by the distance between points along this circle or by the angle between the radii joining the points to the origin. The cosine of this angle, which is a measure of similarity, is the Q-correlation between profiles. The corresponding distance measure can be obtained from the sum of squared differences between the elements of standardized profiles. The relationship between them is $D^2 = 2(1-Q)$.

There are several reasons why intercorrelations among profiles are so widely used despite the fact that this index discards information about dissimilarity.

1. Many investigators are more comfortable using correlations because they are accustomed to such a measure and hence feel they can interpret them. This famili-

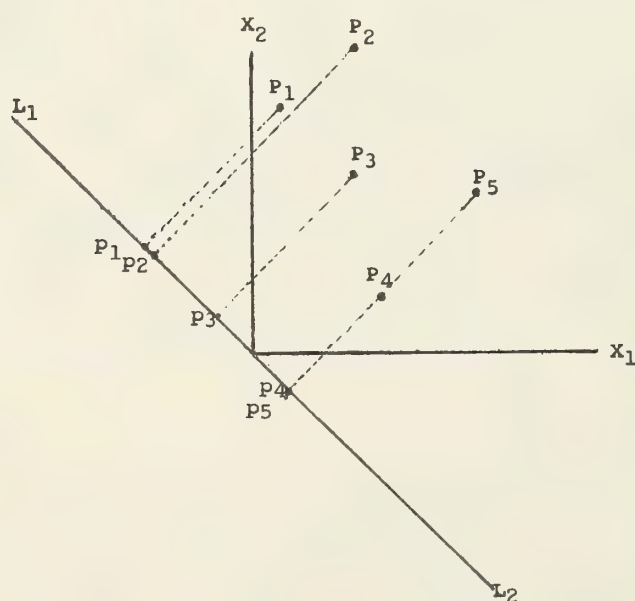


FIGURE 3.—Eliminating dissimilarity attributable to profile elevation.

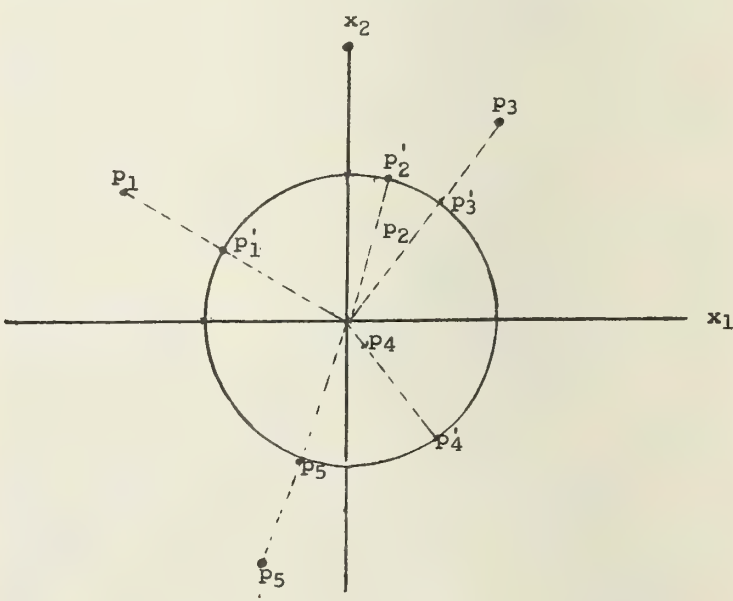


FIGURE 4.—Eliminating dissimilarity attributable to scatter.

arity is in actuality a drawback since it leads to statements about significance levels which are only appropriate under a random sampling assumption about the paired elements. There is no reason to assume, in general, that Q correlations will be distributed like r nor that the expected correlation between persons in a sample is zero.

2. As Nunnally (9) points out, correlations can be handled by powerful methods such as factor analysis which cannot be used on distance measures. The fact that factor analysis is applicable to any matrix of cross-products is not so well known. But while Q correlations can be factored it is doubtful whether the factor analytic model used in this manner is appropriate for grouping persons (17). It is hard to conceive of what is meant by factor loadings of persons on types.

3. Persons attempting to use the distance measure to cluster persons on real data have apparently found that many profiles are drawn into one single large cluster while more extreme or distant profiles do not cluster, but tend to remain isolated from each other (8). Of course, this is exactly what one might expect would happen if the sample is actually homogeneous. Under such circumstances, using correlations because they cluster more readily, seems to be a rather arbitrary solution unless it can be shown that the resulting groupings are more useful or meaningful than the original multivariate distribution of scores.

Normalized vector product.—One additional index of similarity, having the properties of a correlation coefficient but equating profiles which differ only by a constant of proportionality, is the cross-product of normalized profile elements. In terms of our geometric model, this index removes the contribution to dissimilarity of distance from the origin, so that all points are projected onto a hypersphere. Points lying on the same vector are thus treated as coincident while points on closely adjacent vectors are similar regardless of the distance of the

point from the origin. The coefficient which achieves this grouping is

$$\cos \theta_{pq} = \frac{\sum_i X_{pi} X_{qi}}{\sqrt{\sum_i X_{pi}^2 \sum_i X_{qi}^2}}$$

where θ_{pq} is the angle between the vectors from the origin to the points p and q . If raw scores are all positive the coefficient varies from 0 to 1 with "1" indicating that the two points lie on the same vector while "0" indicates the vectors are perpendicular. This index is not invariant for shifts in the origin and hence is useful only when the "zero" on all scales can be meaningfully interpreted. In 1961 (4) I suggested the possible merit of this measure for psychiatric classification based on type and extent of abnormal behavior or symptomatology. For this purpose the origin has a meaning corresponding to "no symptomatology." Overall has made independent mention of this measure (11) and indicated its relation to the distance measure in $k-1$ space. In a more recent article (12) he reports an interesting application of the normalized cross-product in the objective assignment of patients to presently used diagnostic categories on the basis of the similarity of their profiles on a set of rating scales to the typical profile for that diagnosis as judged by a group of psychiatrists.

It should be evident by now that what constitutes similarity or dissimilarity between profiles is primarily a matter of definition. The question is meaningful only with reference to a specific set of variables or attributes scored in a particular way. Even after the elements or variables which constitute the profile have been decided upon, dissimilarities along certain dimensions of the total domain can be ignored or minimized by an appropriate choice of index. Shape, scatter, and elevation can be treated as three separate components of a profile, each of which can be differentially weighted to arrive at an overall index of similarity. Weights can be determined

empirically to maximize predictability for specific criteria. However, if weighting is used for this purpose it is likely that a linear combination of some or all of the variables making up the original profile would result in at least as accurate a prediction with much less work.

The need for new methods of classification of psychiatric illness is often supported by the fact that present diagnostic categories do not relate well to outcome with various therapies. One might ask, however, whether any classification based on clustering is likely to serve this or any other specific purpose optimally. A much better approach to this problem would be to relate the multivariate information contained in some of the presently obtained rating scales directly to outcome under various therapies and then to partition the domain according to boundaries where the various outcome surfaces for different treatments intersect (3). A patient would then be assigned to the treatment which was optimal for the region containing his score set. Classification systems derived in this manner would have a clear meaning and purpose and might also have a better likelihood of relating to etiological factors as well as outcome.

REFERENCES

- (1) CATTELL, R. B. " r_p and Other Coefficients of Pattern Similarity." *Psychometrika*, 1949, 14: 279-298.
- (2) CRONBACH, L. J., and GLESER, GOLDINE, C. "Assessing Similarity Between Profiles." *Psychol. Bull.*, 1953, 50: 456-473.
- (3) CRONBACH, L. J., and GLESER, GOLDINE, C. "Psychological Tests and Personnel Decisions" (2d ed.) Univ. of Ill. Press, Urbana, 1965.
- (4) GLESER, GOLDINE, C. "Logical Constructs Underlying Various Methods of Grouping Persons." In: *Transactions of the Sixth Research Conference on Cooperative Chemotherapy Studies in Psychiatry and Research Approaches to Mental Illness*. Veterans Administration, Washington, D.C., 1961.
- (5) HEERMAN, EMIL F. "Comments on Overall's Multivariate Methods for Profile Analysis." *Psychol. Bull.*, 1965, 63: 128.
- (6) LAZARFELD, P. F. "The Logical and Mathematical Foundation of Latent Structure Analysis."

In: S. A. STOUFFER, L. GUTTMAN, E. A. SUCHMAN, P. F. LAZARFELD, SHIRLEY A. STAR, and J. A. CLAWSEN, "Measurement and Prediction," 1950, Princeton: Princeton Univ. Press.

(7) MAHALANOBIS, P. C. "On the Generalized Distance in Statistics." *Proc. Nat. Inst. Sci., India*, 1936, 12: 48-58.

(8) MCNAIR, D. M., and LORR, M. "Differential Typing of Psychiatric Outpatients." *Psychol. Rec.*, 1965, 15: 33-41.

(9) NUNALLY, J. "The Analysis of Profile Data." *Psychol. Bull.*, 1962, 59: 311-319.

(10) OSGOOD, C. E., and SUCI, G. "A Measure of Relation Determined by Both Mean Difference and Profile Information." *Psychol. Bull.*, 1962, 49: 251-262.

(11) OVERALL, J. E. "Note on Multivariate Methods for Profile Analysis." *Psychol. Bull.*, 1964, 61: 195-198.

(12) OVERALL, J. E. "Application of Classification Procedures in Drug Research." In symposium: "Explorations in Typology with Special Reference to Psychotics," presented at A.P.A., Sept. 4, 1964. Human Ecology Fund, Inc., New York.

(13) REES, L. "Constitutional Factors and Abnormal Behavior." In: Eysenck, H. J. (editor), "Handbook of Abnormal Psychology," Basic Books, Inc., New York, 1961, 344-392.

(14) ROGERS, D. J., and TANIMOTO, T. T. "A Computer Program for Classifying Plants." *Science*, 1960, 132: 1115-1118.

(15) STEPHENSON, W. "Correlating Persons Instead of Tests." *Charact. & Pers.*, 1935, 4: 17-24.

(16) WARD, J. H., Jr. "Hierarchical Grouping to Optimize an Objective Function." *J. Amer. Stat. Assoc.*, 1963, 58: 236-244.

(17) ZUBIN, J. "Taxonomy in the Mental Disorders: A Historical Perspective." In symposium: "Explorations in Typology with Special Reference to Psychotics," presented at A.P.A., Sept. 4, 1964. Human Ecology Fund, Inc., New York.

Supplementary Comments by Dr. Gleser as Presented at the Conference

I had better clarify one thing, and that is we did not develop the measure D or D square for the purpose of clustering, but rather as a means of interpreting profile similarity measures in general.

In the following sessions we will have the opportunity to consider some of the empiri-

cal approaches which have been used in preliminary attempts to develop typologies of psychiatric patients. These methods have in common the fact that they sort out clusters of patients who are maximally similar to each other on the set of measures being used, and minimally similar to patients in other clusters. However, the index which is used to define similarity varies from one study to another. Different indices define somewhat different configurations of scores as maximally similar or equivalent, even when the same data is used.

In this paper I have made an analysis of these similarity indices by placing them in a geometric framework in which they can be contrasted and compared.

Much of the material presented here has been published previously in an article in the "Psychological Bulletin," 1953, written by Lee J. Cronbach and myself. My emphasis at this time is on the following points:

"Similarity" is a construct which is at the basis of all measurement—particularly the measurement of individual differences. When we place objects in one-to-one correspondence with a set of scale points we are implying that objects assigned to the same point are alike in some respect while those assigned to adjacent points are similar. If the scale is ordinal, persons in adjacent categories are more similar than persons in categories further removed. In an interval scale the arithmetic difference between numbers assigned to categories represents the extent to which persons are viewed as dissimilar. It is important to note that the extent of similarity or dissimilarity is always indicated with respect to some characteristic, trait, or dimension. Persons falling in even the two most extreme categories of some scales can be seen as similar or alike in terms of another construct.

The second point that I have made is that there is no one true or most meaningful way to classify people; nor will people alike with respect to one set of attributes or scales

necessarily be more similar to each other than to persons in general on other characteristics. People can be classified into types in as many ways as we can specify the behavioral basis or the physical, mental or emotional characteristics for making the judgment of similarity. Each such classification will probably relate to some aspects of the person's past, present, or future behavior and not to other aspects. Any particular classification will be meaningful only to the extent that it is based on variables or attributes which are related to the broader class of behavior one is trying to predict or control. In psychiatry it is very likely that many classifications will be needed. Thus we would have multiple measures on each person, but now these would be of a categorical, unordered nature, that might be less amenable to powerful methods of statistical analysis.

Thirdly, I have tried to emphasize that all methods of grouping persons into types discards information contained in the original observation or measurement data. Loss of information comes from the coarseness of grouping and also from the index used to assess similarity.

Now, this fact in itself need not be disturbing. All measurement of behavioral phenomena discards information at various stages in the process from observation to final score. The important question is whether the information discarded by the typology is relevant to the purposes of the investigator or whether it is expendable.

In order to answer this question the investigator must have criteria of relevance either in terms of his theoretical model or of the practical purposes for which the classification is intended. The problem we face in psychiatric classification is that there is no agreement at present on a theoretical model, and the purposes for which classification is desired are multiple and most likely related to somewhat different information. Thus, premature typology may well result in obscuring important relationships.

If what is desired at present is a classification system which will relate optimally to outcome under various therapies, a much better approach than clustering is available. The multivariate information contained in presently used rating scales can be related directly to outcome with each specific therapy to determine a best-fitting outcome surface. The intersection of such surfaces for different therapies would provide boundaries for classes within the multivariate domain of rating scales, each class being defined by an optimal therapy. Patients would be classified according to the region containing their score set and given the treatment which is optimal for that region. A classification system derived in this manner would have a clear meaning and purpose.

Multidimensional Representation of Similarity Structures

Warren S. Torgerson, Ph. D.¹

Our major task in this session is to consider some of the current issues involved in the methodology and statistics of classification in psychiatry and psychopathology, in particular, in classification as it pertains to varieties of mental disorder. There seems to be widespread agreement that the present set of diagnostic categories is inadequate, and that the lack of an adequate set of categories is a major limiting factor in studies dealing with causes and cures of mental disturbance.

Classification in the field of mental disorder thus seems to refer to the sorting of individuals into mutually exclusive, qualitatively different classes or types. But it seems to me that this is too narrow an

interpretation. The overall problem is really one of developing an efficient and fruitful means for representing mental disorder. A valid and meaningful structure of mental disorder is needed, and, if that structure turns out to be a set of mutually exclusive classes of individuals, then well and good, a typology of individuals is needed. But if it does not, then such a typology will not serve the purpose.

The overall task of establishing a general structure or framework for representing mental disorder is an immense one. Clearly no single study, no one type of data, no one method of analysis can be sufficient. The major difficulty results from the fact that the research must begin at the beginning, and, at the beginning, objective criteria are not available. The process of establishing an appropriate structure thus almost necessarily implies long-range, bootstrapping operations involving many kinds of data and types of analysis.

The particular corner of the general process with which I will be concerned deals with two different kinds of data. First, the $N \times N$ square symmetric matrix containing as elements indices of degree of similarity with respect to the domain of interest of all pairs of the N individuals. The observations may either be obtained from direct subjective judgments or might be derived from some other procedure—the methods to be discussed make no requirements about how the observations were obtained; they only require that it be reasonable to interpret the observations as monotonically adequate indices of degree of similarity. Clearly the results—and their interpretation—will depend very much on the nature of the observations, but the methods themselves are impartial. Second, and perhaps more central to the present issue, the $N \times n$ matrix of scores or values of N individuals on n variables. Again, the observations might be based on such things as subjective ratings on different attributes, scores on different psychological tests, or objective measures of

¹ Department of Psychology, Johns Hopkins University, Baltimore, Md.

different physiological or behavior variables—the methods to be discussed neither know nor care. Given either type of matrix, the problem is taken to be that of uncovering any meaningful, worthwhile structure that is present, though buried, in the raw data.

Let me digress a moment to dispose of a number of similar problems which I shall subsequently ignore. First, I shall not be interested in simplification for its own sake, as for example, when a continuous distribution is divided into a top half and bottom half in order to simplify or quicken data analysis, or when a structure is deliberately oversimplified in order to make the results for a particular experiment easier to describe. Second, I shall not be concerned with the problems or methods of differential prediction. These procedures can be used to establish classes or categories of individuals. Although the classes so established may have great practical value, the utility of the resulting classification for theoretical purposes may not be great. Third, I shall not be concerned with the problem of determining membership in classes or categories that have already been established. This important problem is being discussed by Professor Rao.

We are left, then, with the basic, fundamental problem: Can we find in the data themselves a basis for establishing a meaningful structure for representing mental disorder. Obviously, if the required information is not present in the data no method of analysis can reveal it. A method cannot uncover a meaningful structure that is not there. It is equally true, however, that an inappropriate method can succeed in obscuring or losing meaningful structures that are present, and this seems to me to be a serious danger with many of the presently available procedures.

A Typology of Present Methodology

A perhaps inadequate—certainly incomplete—review of the recent methodological literature has led me to the following

typology: The methodologists in this area fall into two groups, the typologists and the dimensionalists. This typology seems to have most of the advantages and disadvantages of typologies in general. That is, it seems to make a certain amount of sense, the types obtained are not really new, and it does not really work as well as one would like. Members within a particular type differ markedly from each other; certain members of different types seem quite similar; any single sorting probably lacks reliability and generality; and the prognosis of success of any particular study seems relatively independent of class membership. Nevertheless, the use of types does simplify discussion considerably. One can now discuss the approach and procedures of the typical typologist and dimensionalist and feel free to ignore the fact that most are atypical.

The typical typologist

We take as given either the $N \times N$ matrix of interperson similarities or the $N \times n$ matrix of scores of N persons on n relevant variables. The similarity matrix can be analyzed directly, but the person by variable matrix needs preliminary processing to turn it into an $N \times N$ similarity (or difference) matrix before the standard procedures are applicable. In order to accomplish this transformation, one or more of the many possible indices of pattern similarity is computed for each pair of individuals. (Professor Gleser's paper for this session presents a detailed discussion of a large set of these indices.)

Given an $N \times N$ matrix of interperson similarities (either directly observed or computed from patterns) the next step is to use one (or more) of the many cluster analysis routines to establish clusters or types of individuals (Dr. Lyerly discusses the logic and methods of cluster analysis for this session). Use of the typical clustering procedure results in the assignment of each individual to one and only one class.

The final problem is to apply procedures that enable the experimenter to assign a new individual to one of the empirically established classes. Again, several procedures are available (Professor Rao will discuss them). The typical procedure is a routine that enables the typologist to assign each new individual to one and only one of the established classes.

The typical dimensionalist

Again, we take as given either the $N \times N$ matrix of interperson similarities or the $N \times n$ matrix of scores of N persons on n relevant variables. Although I know of no typical dimensionalist in this context who has worked with the directly observed, interperson similarity matrix, he would surely apply one of the standard routines of multidimensional scaling² to the data. In these procedures, individuals are treated as points on a space of unknown dimensionality. The analytical problem is to determine the dimensionality of the space and the locations of the points (projections on each of the axes or dimensions) in the space. If the space is Euclidean, the axes can be rotated and the final rotated dimensions interpreted. In the final solution, distances between points in the space correspond to some monotonic function of the original observations on the similarity of the corresponding subjects.

Description of the treatment of the $N \times n$, person by variable matrix by the typical dimensionalist is more complicated. Only one thing is certain: Sooner or later at least one factor analysis will be performed. One strategy is to obtain one of several cross-product indices over individuals, and then to factor the resulting $n \times n$ matrix of cross products of the variables to obtain a smaller set of more basic traits or dimen-

sions for describing the differentiating characteristics of the subjects. Another strategy is to obtain one of the several cross-product indices over variables and then to factor the resulting $N \times N$ matrix of cross products of the individuals, again to obtain a smaller set of more basic dimensions for describing the differentiating characteristics of the subjects.

A third strategy would be to do both: An initial factor analysis of the variables to obtain a new set of variables that are relatively independent or uncorrelated with each other and then a factor analysis of the individuals to obtain the final configuration of people. Again, subjects are described by their projections (or scores) on the different dimensions of the space.

Discussion

I would like to call into question the approaches of both my typical typologist and my typical dimensionalist (with some justification, the reader might wish to substitute stereotypes or straw men). Let us leave aside, at least for the moment, such secondary, though analytically extremely important, questions as—

- (a) whether or not the variables of the $N \times n$ matrix ought to be left as observed, or centered, or standardized;
- (b) whether the index of pattern similarity should be the correlation coefficient, covariance, cross product, normalized cross product, or Euclidean distance or distance squared, and whether these indices should be based on correlated or uncorrelated variables; and
- (c) which clustering technique or factor analysis procedure ought to be used.

Rather, let us concern ourselves with the implicit model underlying each approach, the requirements each places on the data, and the outcome or results obtained by each when it is used.

² Kruskal, J. B. "Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis." *Psychometrika*, 1964, 29, 1-28. Shepard, R. N. "The analysis of proximities: Multidimensional scaling with an unknown distance function." *Psychometrika*, 1962, 27, 125-140, 219-246. Torgerson, W. S. "Theory and methods of scaling." New York: Wiley, 1958.

The typical typological approach

The typical cluster analysis procedure seems based on the notion that the underlying structure is one of discrete, qualitatively different classes of individuals. The underlying model or theory is either simply that of an unordered set of discrete nominal classes, or for some clustering procedures, that of a hierarchical set of discrete classes with each class divided into subclasses.

The outcome of a cluster analysis is the assignment of each individual in the sample to one and only one class (at a given level). A second implicit assumption thus seems to be that nature permits one and only one mental disorder per person. Persons with similar patterns of scores over all of the variables are grouped into one class; those with dissimilar patterns are sorted into different classes.

Although my own training and experience in the area of mental disorder is minimal, yet it seems to me that the requirements of the implicit model underlying cluster analysis are much too strong. First, even assuming that there are such things as discrete, qualitatively different mental disorders, an assumption that nature permits only one per person seems reckless. Even an assumption of independence in distribution of mental disorders would seem risky. Certainly physical disorders do not behave this way.

Second, even if we accept the notions that there are discrete, qualitatively different, mental disorders and that they are distributed one per person, it still does not seem likely that subjects having a particular disorder would be homogeneous with respect to all variables in the set. Some of the variables might well be at least partially sensitive to such things as overall degree of disturbance regardless of the particular disorder, or to personality, behavioral, or physical traits not relevant to the particular set of disorders involved.

Since clustering is based on indices of overall pattern similarity, the presence of such variables can easily lead to a classification of individuals that is unrelated or only partially related to the given set of discrete mental disorders.

The typical dimensional approach

In this approach, the underlying model would seem to imply that the structure of mental disorder is one of continuous, quantitative differences of degree. Individuals are represented by their values or scores on an underlying set of quantitative traits or dimensions.

It is clear that a dimensional interpretation is always possible. Any structure whatsoever underlying the data can be embedded in a Euclidean space. Given the $N \times n$ matrix of scores of persons on variables, one need only consider the variables themselves as axes or dimensions of a Euclidean space, and the observed scores of individuals as projections on these axes. Such a representation obviously preserves all of the information in the original matrix. Furthermore, if the individuals all lie within an r -dimension subspace of the total n -dimensional space defined by the n variables, then a factor analysis of raw cross products between individuals will yield an r -dimensional configuration of individuals that also preserves all of the original information. A skilled factor analyst can ordinarily find a rotation of the axes that will enable him to interpret each of the important dimensions. The disorder associated with a given individual can therefore be described by his loadings or scores on the final set of dimensions or traits.

But even though a dimensional representation can always be made, it need not be the most meaningful representation. The dimensional representation can easily lead one to overlook special features of the configuration; features which, if noticed,

would lead to a vastly different interpretation of the underlying structure. Suppose there really are independent, discrete, qualitatively different mental disorders. Will the dimensions obtained correspond to the disorders in a one-to-one fashion? Not in general. As we shall see, under many circumstances even the number of independent, qualitatively distinct disorders will not equal the number of dimensions.

The dimensional approach based on the factor analysis of individuals is further hampered by the sensitivity of the obtained configuration to the particular cross-product index chosen by the experimenter. Four such indices are in use: the cross product itself, the normalized cross product, the covariance, and the correlation coefficient. The cross product preserves all of the information. The normalized cross product simply stretches or shrinks the vectors of the cross-product solution so that they are all of unit length. The covariance and correlation are similarly related. These latter two coefficients, however, remove or eliminate effects due to differences in overall level between the individual profiles—a procedure that under many circumstances seems questionable. Many observable variables of interest are bipolar, i.e., they could be originally scored in either direction. Both covariance and correlation are highly sensitive to the direction in which such variables are scored.

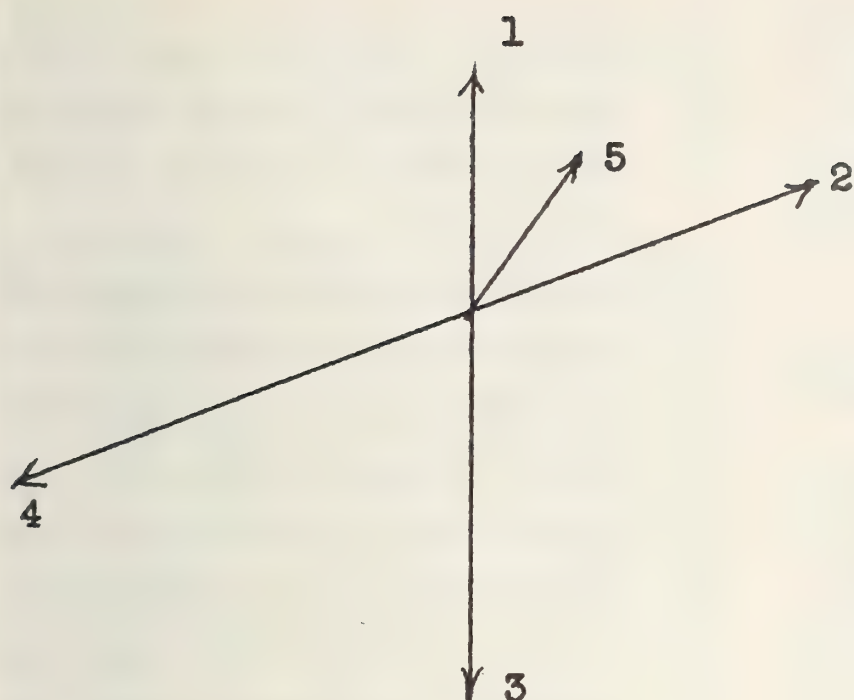
Use of one coefficient rather than another markedly changes the appearance of the final configuration. As a result, the final rotation and interpretation of underlying traits cannot be expected to remain invariant over changes in coefficient. One small, error-free example of the effect of choice of coefficient is shown in figure 1. The configurations show the results obtained when each of the four coefficients is used on the following matrix of scores of five hypothetical individuals on three hypothetical variables:

Individuals	Variables		
	1	2	3
1	$\sqrt{2}$	0	2
2	$\sqrt{4.5}$	3	0
3	$-\sqrt{4.5}$	0	-3
4	$-\sqrt{8}$	-4	0
5	$\sqrt{2}$	1	1

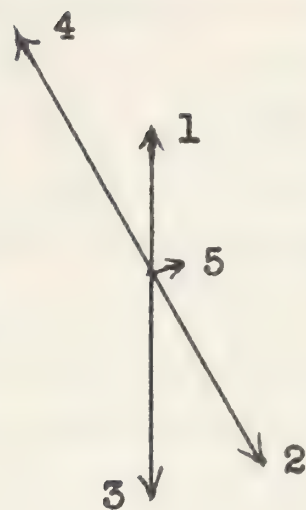
In this particular example, although all configurations are two-dimensional, their shapes differ markedly. If the same exercise were repeated with variable 1 omitted, the differences would be even more striking: the cross product and normalized cross-product configurations would change only slightly, but the covariance and correlation configurations would be reduced to a single dimension. It might also be noted that clustering procedures here would fare no better: different coefficients would lead to different clusters.

It would thus seem that, if the task is taken to be the uncovering of meaningful structure in the data, both approaches have various weaknesses. The clustering procedures of the typological approach force individuals into mutually exclusive classes that are either nominally or hierarchically structured, whether or not such classes exist in any meaningful sense. The dimensional approach, as it is ordinarily used, forces an exclusively dimensional interpretation of the underlying structure. But purely dimensional and purely exclusive-class structures are only two of the many possible structures that might serve as a meaningful basis for representing mental disorder. One of the more attractive alternatives, for example, would involve a mixture of the two: a set of classes (either exclusive or nonexclusive) representing specific disorders, with one or more quantitative dimensions superimposed.

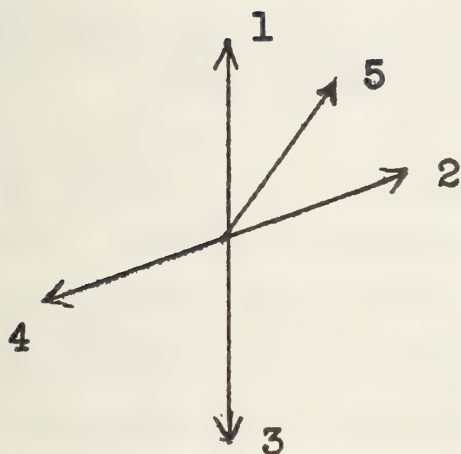
It would seem that the problem really requires a method of data reduction or analysis that can enable the experimenter to determine the appropriate structure, rather



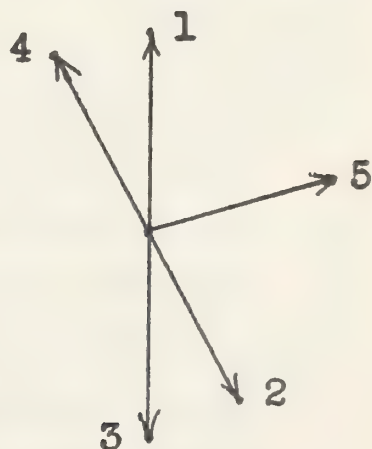
1A. Cross Product



1B. Covariance



1C. Normalized Cross Product



1D. Correlation

FIGURE 1.—Effect of coefficient on configuration.

than procedures or approaches that impose one or another structure on the data. I am beginning to believe that methods which have been available for quite some time can serve the purpose if they are used insightfully. The methods are those of multidimensional scaling and factor analysis. I should emphasize strongly that I refer to the methods only as data reduction procedures, and not as psychological theories, approaches, or models. The theoretical side of

factor analysis does carry with it implications of continuity and of quantitative dimensions, primary factors, and the like. Similarly, the multidimensional scaling approach easily leads to expectations of configurations and invariances which an underlying structure can rule out. But the methods as methods actually carry no such implications or expectations.

We have seen that any structure whatsoever can be embedded in a space of ap-

appropriate dimensionality. Differences between structures that are logically distinct can therefore only be revealed by the restrictions each structure makes on the shape or character of the configurations of points it can generate. My conjecture is that the experimenter can use the information contained in the shape of the obtained configuration to infer the nature of the underlying structure.

In the section to follow, a few examples of such configurations will be presented. But since configurational shapes in factor analysis depend very much upon, first, whatever preprocessing is performed on the n variables and, second, which index of association is used, it follows that any examination of the configuration restrictions placed on the data by different underlying structures requires first a precise specification of the analytical procedures involved. My discussion will be based on the following two analytical procedures:

Factor analysis

- (1) It is assumed that the $N \times n$ matrix of observations is scaled so that the sum of the scores over subjects for each variable is zero. Ordinarily the variables would also be scaled so that their variances are all equal also, but most of the conclusions to be drawn do not depend upon equating of variances.
- (2) Raw cross products are computed for all pairs of subjects, giving an $N \times N$ symmetric matrix with cross products in the off-diagonal cells and sums of squares of scores in the diagonal cells.
- (3) Any standard factor analysis routine that yields projections on an orthogonal set of axes is then used on the cross-product matrix to obtain the configuration of individuals.

Multidimensional scaling

- (1) For a directly observed $N \times N$ matrix of interperson distances: The dis-

tances are converted to scalar products from an origin at the centroid of the individuals and the matrix of scalar products is factored as indicated above.

- (2) For an $N \times n$ matrix of individuals by variables: Euclidean distances are computed between individuals (considering the n variables as n orthogonal axes) the distances then converted to scalar products from an origin at the centroid of the individuals, and the scalar-product matrix factored as before. It is interesting to note that, under these circumstances, it turns out that the scalar-product matrix is precisely the cross-product matrix of the factor analysis procedure. Therefore, under the stated circumstances, the multidimensional scaling and factor analysis procedures are equivalent.

Configurational Restrictions Imposed by Some Simple Structures

First, let us consider the simplest of all structures: the structure of independent, qualitatively different, mutually exclusive classes. Each individual is thus a member of one and only one class. Members of the same class are similar and members of different classes are dissimilar. With k classes and error-free data, the resulting configuration is a set of k distinct points located in the corners of a $k-1$ dimensional simplex. For example, if there are just three classes, the individuals are located at one or another of the three vertices of a triangle in a two-dimensional space, as illustrated below:

○

○

○

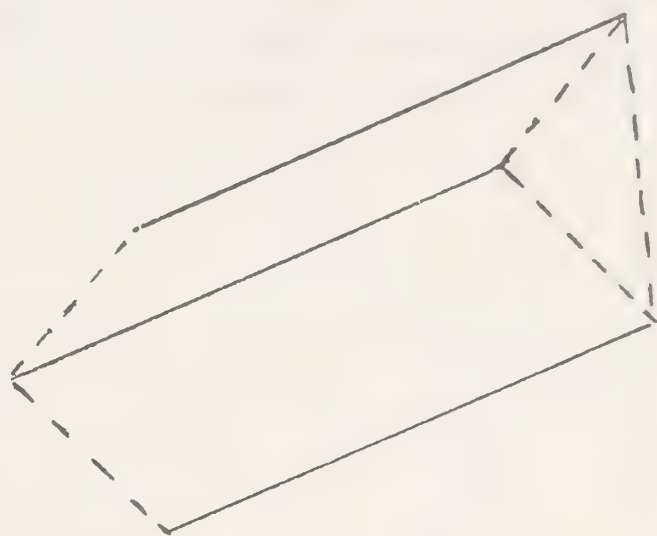
Addition of random error would simply change the three points to three clusters of points in the two-dimensional space, and add $n-3$ or $N-3$ (whichever is smaller) small error dimensions.

Since only two dimensions are needed to represent three classes, it could be argued that a dimensional interpretation would be more parsimonious than a class interpretation. But this argument would ignore the fact that the dimensional interpretation tells us nothing about the most striking feature of the configuration, namely, that individuals can occur at only the three locations.

Cluster analysis of the data would also yield the appropriate number of classes, of course, since the nominal, exclusive-class structure is the model underlying most clustering procedures.

In the structure just considered, degree of similarity was assumed to be entirely due to class membership. Suppose now that degree of similarity is determined only in part by class membership and also in part by degree of difference on one or more quantitative dimensions that cut across class boundaries. This would occur, for example, if some of the variables were sensitive to overall degree of disturbance, regardless of the type of disturbance involved. Specifically, we consider the case where the squared Euclidean distances between individuals consist of two additive components: one due to class membership and the other due to differences on quantitative dimensions. Given k classes and d orthogonal, quantitative dimensions, the configuration would consist of k , d -dimensional hyperplanes located in a total space of $d+k-1$ dimensions. The projections of the individual points onto the $k-1$ dimensional subspace determined by class membership would cluster at the vertices of a $k-1$ dimensional simplex as before. Projections of points onto the d quantitative dimensions would have no special distinguishing characteristics. For three exclusive classes and one orthogonal dimension, the points would

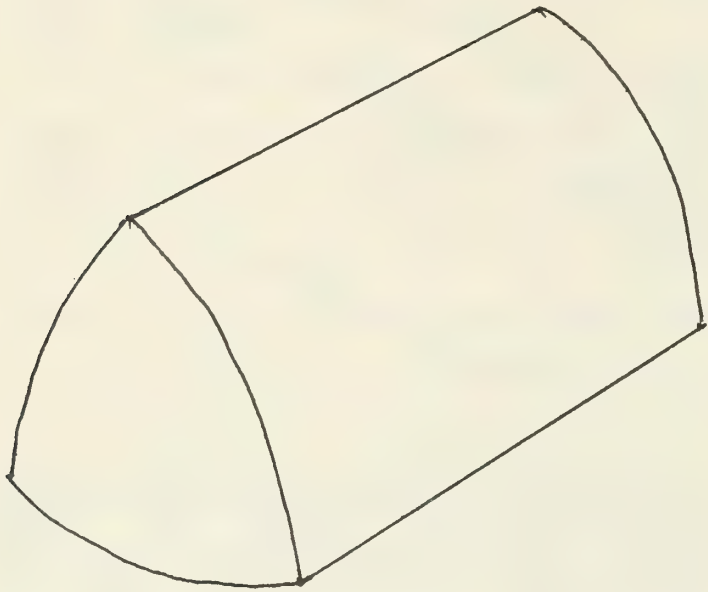
therefore be located in a three-space, but only on the three lines corresponding to the edges of a triangular prism, as shown by the solid lines in the diagram below:



Application of cluster analysis procedures to data based on a structure such as this would be unrevealing. Any clusters obtained would likely represent an unknown mixture of class and dimensional variables. The usual factor interpretation would also be misleading, since the analyst would attempt to interpret the k discrete classes in terms of $k-1$ continuous dimensions or traits.

As a final illustrative structural model, let us consider the case where each individual's disturbance is composed to differing extents of any or all of a set of discrete, qualitatively different disorders. Any individual might therefore be a member of several different classes, and to different degrees. Furthermore, let us assume as before the presence of overall quantitative dimensions (e.g., overall seriousness) that are independent of the particular composition of disorders characterizing a given person. Since this structure is a little more complex, an analogy from the physical world might be helpful. Consider the alloys of copper, tin, and zinc, and let us take as our sample a number of specimens (individuals) varying in size (overall seriousness) and composed of differing proportions of the three metallic elements (disorders). Assume that the observational procedure or set

of variables is sensitive to both size and composition. The precise configuration obtained from such a structure will clearly depend upon the specific way in which this sensitivity occurs. Under many reasonable circumstances, however, it would seem that the following configuration structure would be closely approximated:



All specimens would lie on or within the boundaries of the curved prism shown. Specimens composed of a single metallic element would lie on one of the three edges of the prism, with the location along the edge dependent on the size of the specimen. Those composed of two metallic elements would lie on one of the three outside surfaces. And those composed of all three metals would lie within the interior of the prism. Finally, all specimens of a given size would lie within that closed triangular section of the prism that corresponds to their size.

Again, neither the purely typological approach nor the purely dimensional approach would be likely to reveal the nature of the underlying structure. Both could yield answers, of course, since both can always yield answers. But neither would indicate that the specimens vary in size and are composed of various proportions of three elements.

The structures just described are only a small sample of those which might be found to underlie mental disorder. But it seems

likely that other reasonable structures would also generate configurations as recognizable as these. If so, then the combination of dimensional method (factor analysis or multidimensional scaling) with a structural interpretation of the resulting configuration can yield results of both theoretical and practical utility. The approach could indeed be used to determine an underlying structure rather than to impose a given structure on the data. If my conjecture is correct, it would seem that structural interpretation of dimensional configurations could do much toward establishing a meaningful framework for representing mental disorder.

A Survey of Some Empirical Clustering Procedures

Samuel B. Lysterly, Ph. D.¹

Introduction

One of the very distinguished participants in this conference has written: "No formal rules can be laid down for finding * * * clusters because a cluster is not a well-defined term" (Rao, 1952, p. 362). Yet the notion persists that there can be, or may be, or should be, ways of assigning people to groups in such a way that knowing a person's diagnosis or classification will significantly aid professional workers in helping him in medical, vocational, educational, or other situations. If we did not believe this, we (or many of us) would not be here today.

Dr. Greenhouse has defined the clustering problem as distinguished from the separate but related problems of discrimination and classification. It is my purpose to present accounts of some of the empirical

¹ Psychological Assessment Associates and Bureau of Social Science Research, Inc., Washington, D.C.

attempts that have been made in recent years, most of them by social scientists, and to call attention along the way to some of their characteristics and some of the difficulties in the several approaches.

The clustering problem begins with a matrix of numbers representing comparisons of every person in a sample with every other person. For n individuals we have a table of $n \times n$ such numbers. They may be indices of similarity (e.g., correlation coefficients or covariances) or dissimilarity (e.g., measures of "distance"). The decision concerning which kind of index to use will ordinarily be determined by the purpose of the investigator and the nature of the data. Dr. Gleser in her paper has discussed several of these alternatives and their major implications and I shall not comment further except to underscore the warning she has given more than once that the structure resulting from any attempt at empirical clustering depends heavily upon the choice of a similarity (or dissimilarity) index (Cronbach and Gleser, 1953; Gleser, 1961).

With a matrix of interperson similarities or dissimilarities how can an investigator best proceed to detect clusters, if clusters exist? In general, workers have attempted to form groups such that each member is closer to other members of his group than he is to members of other groups. This criterion, here loosely stated, permits of various interpretations. The requirements of the problem, the characteristics of the population, and the choice (or availability) of appropriate variables may very largely determine whether a clear-cut galactic separation of groups is possible, whether a "lumpy" texture is found, or whether several areas of doubtful concentration in the n -dimensional space can be identified, leaving many individuals to be unclassified "loners."

Early attempts at forming clusters, one notable example being the B-coefficient (or "coefficient of belonging") of Holzinger and Harman (1941), were directed toward the partitioning of a correlation matrix into

smaller matrices to facilitate subsequent factor-analytic or other procedures. The B-coefficient was designed to place into subgroups those variables whose average correlations with each other were higher than their correlations with variables in other subgroups. Although it was not intended for use with interpersonal measures, a number of investigators have used the B-coefficient to classify individuals, at least tentatively. It is defined as "100 times the ratio of the average of the intercorrelations of a subset or group of variables to their average correlation with all remaining variables." This criterion is admittedly a rule-of-thumb, even in the situation for which it was originally designed, and has no particularly useful statistical properties.

Perhaps most of the empirical clustering methods in the social and psychological sciences have taken the general approach implied by the B-coefficient. The matrix of interperson indices is inspected and the two closest individuals are selected to form the nucleus for a type or group or cluster. Then other individuals or pairs are examined to become nuclei of new clusters or to be added to existing clusters. Various sequences or rules have been followed, and various criteria for inclusion, exclusion, or reassignment, some planned objectively (and hence adaptable to computer methods) and others dependent upon the investigator's judgment at various points in the process. Since the rules are essentially arbitrary there are usually a number of individuals in the sample who are left unassigned to groups. Whether this is desirable or not depends, of course, upon the purpose of the investigator.

These methods, which begin with the pairing of similar individuals and then adding to each such nucleus others according to their resemblance to cases in established nuclei, we shall call the "synthetic" approach to the cluster problem.

There is a second group of clustering methods which the statistician may possibly find more compatible with his interests. We will call this the analytic approach.

Instead of beginning with n individuals, each a "cluster of order one," and successively combining pairs and larger groups until all or most have been assigned (depending upon the arbitrary criterion adopted), the investigator begins with the entire sample as one cluster and asks "How can I divide this into *two* groups, each of which is more homogeneous with respect to some criterion or standard than is the total sample *and* more homogeneous 'on the average' than would be the case if *any other* partitioning into two groups were made?" The criterion may be an external one, such as therapeutic outcome, occupational success, etc., or an internal criterion, such as a minimum within-groups sum of squares.

Next, having divided the original sample into two groups, maximizing internal homogeneity (which in effect involves examining each of the possible $(2^{n-1} - 1)$ partitions), the investigator may analyze each of them and search for ways to divide them into further subgroups. This sequence of steps may be continued and the results tested at each stage until the final configuration meets some criterion (e.g., by the discriminant function or within groups/among groups analysis of variance, although an exact test of an appropriate null hypothesis for such a procedure is not known). Many investigators may wish to examine several alternative solutions and select one which makes the most sense for their purposes. This understandable and commendable attitude does, however, further complicate any evaluation of the statistical properties of the outcome.

The result of this series of operations will ordinarily be an hierarchical "tree" configuration of groups, consisting of a "trunk" (the original undifferentiated sample), one or more orders of "limbs" and "branches," and finally the "twigs" (the ultimate smallest groups which cannot be further subdivided). The configuration need not be symmetric. Some ultimate

categories may be at the limb or branch level.

Two characteristics of the analytic approach in comparison with the synthetic are: (1) It is more objective and hence more readily programmed for computers (at least in the forms in which recent investigators have used these methods, though not necessarily in general); and (2) it assures that every individual is assigned to one of the ultimate groups, provided some quasi-statistical criterion is used to terminate the process (such as a predetermined within-groups sum of squares or a minimum number of cases in the ultimate categories).

Obviously, if either the synthetic or the analytic procedure is allowed to proceed unchecked by any rule of "when to stop," the "Sorcerer's Apprentice" will take charge. The synthetic approach will ultimately assign everyone to a single type, and the analytic will finally split the entire group into n classes, each containing one person.

Most of the attempts at empirical clustering have been step-wise and/or iterative procedures. A solution which has some obvious appeal is that the investigator form every possible arrangement and test each such arrangement against the criterion he has chosen. In other words, he would divide his subjects into every possible set of two groups, three groups, etc., and test every such set of partitions.

The difficulty with this approach is that with samples of even moderate size the problem is beyond the ability of even the fastest and most capacious modern computer to handle. The number of ways of classifying n individuals into r groups is $n!/r!$ times the coefficient of x^n in the expansion of the generating function $(e^x - 1)^r$. Table I lists these for samples through size 12. For a sample of 16, which is certainly as small as most investigators would want to use, the total number of arrangements is more than 10 billion! Hence the need for shortcuts, approximations, and

TABLE I.—Number of Ways of Assigning n Persons Into r Groups¹

r													
n	1	2	3	4	5	6	7	8	9	10	11	12	Total
1	1												1
2	1	1											2
3	1	3	1										5
4	1	7	6	1									15
5	1	15	25	10	1								52
6	1	31	90	65	15	1							203
7	1	63	301	350	140	21	1						877
8	1	127	966	1,701	1,050	266	28	1					4,130
9	1	255	3,025	7,770	6,951	2,646	462	36	1				21,147
10	1	511	9,330	34,105	42,525	22,827	5,880	750	45	1			115,975
11	1	1,023	28,501	145,750	246,730	179,487	63,987	11,880	1,155	55	1		678,570
12	1	2,047	86,526	611,501	1,379,400	1,323,652	627,396	159,027	22,275	1,705	66	1	4,213,597

¹ For those interested in combinatorial problems, the frequency in any cell in the above table can be calculated by the recursion formula:
 $f_{(n)(r)} = f_{(n-1)(r-1)} + rf_{(n-1)(r)}$.

iterative approaches to the clustering problem.

Recent Studies Using Clustering Methods

The following are summaries from the literature, mostly from the medical and social sciences, of reports of clustering attempts. (There are interesting developments of similar methods in such fields as biological taxonomy, pattern recognition, linguistics, weather forecasting, etc., which some day may be brought together for the benefit of the entire scientific community, provided that the electronic computer continues to develop and improve its skills in communicating between scientists of different disciplines more effectively than the scientists themselves have been able to accomplish.) The major purpose of these notes is to provide conference participants with examples of recent and current efforts in this area. The order of presentation is largely arbitrary, since there was no immediately obvious way of organizing (or clustering) these reports.

In a number of papers L. L. McQuitty has reported the development of a group of related techniques for analyzing similarities between people (or institutions, occupations, etc.) into hierarchical groupings. (See references for a representative list of papers.) In its simplest form the numbers of agreements on a set of dichotomous items for each pair in the sample are recorded and clusters are formed through successive steps of combining pairs with high agreement indices with individuals or with other pairs. Several variations in method, including extension into the use of continuous rather than discrete data, and several criteria for optimum grouping have been proposed for different kinds of problems. McQuitty's methods apparently have not been used in any extensive study of differential psychiatric diagnosis. It would be difficult to apply to problems involving large numbers of items (variables) and subjects. When only

dichotomous responses are used to construct agreement indices which do not take into account the identity or the content of the items involved, it would seem that much useful information is lost and that the interpretation of any dimensional structure may thereby be difficult.

The "latent structure" model of Lazarsfeld (1950; see also Gibson, 1959 and 1961) has been proposed as a method for identifying clusters and has been considered an analogue (or a substitute) for factor analysis for problems involving qualitative or dichotomous data. Certain of its assumptions and the limitations on the kind and amount of data which can be used probably account for the fact that it apparently has not been used in attempts at psychiatric classification.

Gerard and his associates (1964) have published a preliminary report of a large-scale study of schizophrenics and normals in which a large number of psychological, physiological, biochemical, and sociological variables were employed. Several approaches were used, including factor analyses of various subtests of the variables and clustering techniques including a "successive factorial screening" method which reportedly led to the successful classification of 67 out of 76 patients into 7 identified schizophrenic types. I will make no further comment at this time, since we will have a fuller and more recent account of the project later in this conference.

Cattell (1965) has endorsed the use of factor analysis in clustering problems with some interesting qualifications. He says that "* * * categorizing and taxonomy [are] only a secondary or nonessential part of factor analysis * * *" but suggests that it can add to a "causal understanding" of the resulting structure of a cluster analysis. Cattell proposes an iterative procedure using his coefficient r_p (a function of chi-square) rather than the product-moment coefficient, to derive a first approximate structure by factor analysis. (Nunnally (1962) and others have objected that this

coefficient results in a non-Gramian matrix and hence there is no assurance that the standard methods of factor analysis are valid.) After the first clustering, additional trials using the "central type of each subgroup so defined" is expected to lead to a final satisfactory classification. This is called an "iterative type-dimension analysis." It could be lengthy and laborious when applied to a problem of more than moderate size and complexity.

Nunnally (1962) has demonstrated a "raw score" factor analysis using cross-products which are not standardized by adjusting for differences in level or dispersion and shows that such an analysis can preserve the mutual distances and other characteristics of the original data. Overall (1964) has objected that correlations among the variables will violate the assumptions of Euclidian space and insists upon transformations to an orthogonal reference system which will preserve true distances and directions within a standardized set of dimensions. It seems to me that if distinct clusters are present, they should be recognizable regardless of scale differences or considerations of orthogonality. Monotonic transformations of the defining variables or rotations to principal axes would not be expected to clarify or to distort the configurations from a *topological* point of view. And the identification of clusters, whatever their parametric properties, is the first duty of cluster analysis. Whether Nunnally's factor-analytic approach is comparable in efficiency to others must await further comparative trials.

A recent example of the use of factor analysis in a classification procedure is the work of Katz (Katz and Cole, 1963; Katz, 1964). A sample of profiles is factor-analyzed and types are defined by averaging the profiles of several individuals who stand highest on each of the factors found. Then members of the larger pool of patients (not used in the factor analysis) are assigned to that type with whose aver-

age profile the individual's own profile correlates most highly. Arbitrary levels are used to exclude ambiguous cases (cases that do not correlate sufficiently highly with any profile or which correlate too highly with more than one). A method such as this perhaps should not be called "cluster analysis" in the strict sense. It seeks to group together persons who are alike in being relatively high on one *dimension* and relatively low on all others, regardless of whether there are boundaries or discontinuities or whether there are interior clusters of individuals overlooked because they are not extreme in any dimension. Perhaps partly as a consequence of this, a rather large proportion of Katz' subjects could not be assigned to his types.

Lorr and his colleagues prefer to use correlations or similar indices that are not affected by level. They select an arbitrary value of the coefficient (say the 0.02 significance level) and list with each profile the others which meet this criterion, taking as a first "pivot" that profile with the largest number of related profiles. Then a second profile with the largest average correlation with those in the first list is added, then a third with the highest average correlation with profiles associated with the first two. These are now the nucleus of a cluster, and others are added in the order of their average correlation with the cluster members until none remain that meet the criterion. Similar steps are followed to identify additional clusters among the remaining profiles. This procedure, here greatly oversimplified, involves some arbitrary decisions and, as other synthetic methods, may leave many individuals unassigned (41 percent in one study).

Sawrey, Keller, and Conger (1960) have used a method based upon distance functions. Arbitrary limits for similarity and dissimilarity are set. Profiles are successively compared with each other and to the centroids of selected nucleus groups until each has been assigned to a group or has

been judged unassignable. The procedure as described seems reasonably objective, but somewhat cumbersome.

Ward (1963) and Ward and Hook (1963) approach the clustering problem from the hierarchical point of view. In addition, they propose an "objective function" (a functional characteristic of the data which reflects the investigator's purpose) as a criterion to guide the clustering procedure. In their worked example, using the data from the Sawrey, Keller, and Conger study, the criterion is a minimized sum of squares of deviations within groups. They proceed by reducing the number of groups from n (the original number of individuals) to $n - 1$, to $n - 2$, etc., successively combining individuals or subgroups in an order which at each stage seeks to minimize the within-groups (and hence maximize the between-groups) sums of squares. The method is objective and readily adaptable for computer operation. Since it ultimately will reduce the data to a single cluster, the investigator must decide at some stage that his purpose has been accomplished.

Edwards and Cavalli-Sforza (1965) propose an hierarchical method using distance functions that seeks to minimize within-cluster sums of squares of distances. They begin by dividing the sample into two parts by considering all of the $2^{n-1} - 1$ possible splits, then repeating the process with each of the two clusters thus discovered, until a solution is found that does not materially improve the discrimination. (This method does not guarantee, of course, that the final sum of squares will be an absolute minimum at any stage of the clustering.) They have applied this procedure, which apparently can be readily programmed for computers, to Rao's data (1952) on 12 Indian castes and find a grouping only slightly different from that reported by Rao. In their opinion a straightforward analysis of variance of the among-groups and within-groups sums of

squares is a satisfactory test of the efficiency of the resulting clusters.

Forgy (1963) has proposed a method that has considerable intuitive appeal. Individuals are considered as particles in an n -dimensional Euclidian space, attracting each other according to Newton's law. (Apparently at the beginning of the process they must all be stationary, e.g., no orbits having been established, and in a viscous medium to prevent a buildup of momentum, and all of equal mass.) Under these conditions the points would move toward each other to form clusters (if there be potential clusters in the data) in a manner and a sequence completely predictable from the matrix of Euclidian distances. A computer program to accomplish this has been prepared and tried out on some artificial data. If a clustering procedure such as Forgy's is considered to be operating in "real time," a problem arises in that some meaningful and useful clusters, depending upon their sizes and relative positions, may form and coalesce before others elsewhere have yet formed; and larger clusters may engulf neighboring small clusters. Some form of monitoring or surveillance would seem to be necessary. (Of course, unless there is intervention at some point, all will end up in one concentrated mass.) Forgy has recently (1964) reported some results of his procedure in comparison with several others. He considers that none is entirely satisfactory.

Saunders (Saunders and Schucman, 1962; Saunders, 1964), using distance functions, begins by regarding every individual in the sample as a "cluster of order one." Then with each cycle of a comparison process the best addition of one case to each cluster is identified. These searches are made in parallel to achieve computer economy. This then requires a search of the new clusters for duplications after each cycle, and a final listing of "closed clusters" or "nodes" which may then be examined for such of their characteristics as the investigator may be interested in (e.g., mean pro-

files, variance-covariance matrices, further iterations, etc.). At the judgment of the investigator, the procedure may be stopped at any cycle without attempting to continue until every profile has been classified. On the question of the number of clusters that should be determined from a given set of data, Saunders proposes that "the number of syndromes isolated plus the maximum internal dimensionality of any of these syndromes should not be greater than the number of dimensions of input information." This is apparently an attempt to establish a criterion similar to that regarding the number of dimensions identified in a common-factor analysis as a function of the rank of the correlation matrix. I fail to see the usefulness of such a restriction. Factor analysis is concerned with the number of *dimensions* within a multidimensional space. The number of *clusters* may conceivably greatly exceed the number of dimensions.

Thorndike (1953) has attempted to make iterative, partly subjective cluster assignments based upon distances. He begins by taking the two farthest items (in his example jobs as characterized by their required aptitudes) and placing them into different clusters. Then, he begins the third cluster by assigning to it the case farthest from either of the first two. When the required number of "strangers" has been found, he then assigns in turn to each the one from those remaining which is nearest. Routine application of his procedure leads to clusters all of which have an equal number of members, and subsequent adjustments are made to reassign cases closer to members of another cluster than to their own original ones. A considerable amount of trial-and-error is involved, some of which might now be handled by modern computer methods which were not generally available at the time of Thorndike's report on this effort. Thorndike admits some dissatisfaction with his solution, including the pervasive problem of "when to stop" and the questions of what assumptions are realistic

in order to devise a completely objective procedure and appropriate significance tests for a solution. Although his purpose was to study job families rather than clusters of individuals as such, his goals and his problems are basically the same as those encountered in the fields of personality and psychopathology.

Zubin and his associates (1963, 1964) are working with a method which begins with a distance matrix. Like Thorndike they find the two profiles farthest apart, to each of which is added others whose distance is less than some arbitrary value until all profiles have been assigned to one or the other of the two. These two groups are then tested for multivariate homogeneity, and if they do not meet a satisfactory criterion, the procedure is repeated and three, four, etc., groups are sought until a satisfactory structure is found. (The entire sample may be tested for multivariate homogeneity at the outset, and if the null hypothesis is not rejected, no clusters are assumed.) This procedure, which appears to be a descendant of K. Pearson's two-population univariate case (1894), has the considerable disadvantage, it seems to me, of depending upon certain assumptions of normality and upon conditions that many workers will not readily accept. In particular, the normality assumption is a difficult one.

Some Final Remarks

1. We should seek further advice from the mathematicians and statisticians (although it may be that they have contributed about as much as we can expect, considering some of our peculiar requirements).

2. We need more cross-validation, i.e., a clustering method should be used with several sets of data drawn from the same and different populations to test its generality.

3. We need to have the same set of data analyzed by different clustering methods to assess the comparability of the several approaches.

4. We need more experimental studies of the utility of various classification or diagnostic systems for various purposes (e.g., therapeutic outcome).
5. And finally, but perhaps most important of all, we need to marshal all the theoretical insights and experimental findings at our disposal for the challenging and never-ending task of discovering and properly using the most valid and efficient variables for our several important goals.

REFERENCES

CATTELL, R. B. "Factor Analysis: An Introduction to Essentials." "I. The Purpose and Underlying Models." *Biometrics*, 1965, 21, 190-215. (a)

CATTELL, R. B. "Factor Analysis: An Introduction to Essentials." "II. The Role of Factor Analysis in Research." *Biometrics*, 1965, 21, 405-435. (b)

CRONBACH, L. J., and GLESER, G. "Assessing Similarity Between Profiles." *Psychol. Bull.*, 1953, 50, 456-473.

EDWARDS, A. W. F., and CAVALLI-SFORZA, L. L. "A Method for Cluster Analysis." *Biometrics*, 1965, 21, 362-375.

FORGY, EDWARD W. "Detecting 'Natural' Clusters of Individuals." Paper presented at Western Psychol. Assoc. meetings, Apr. 19, 1963.

FORGY, EDWARD W. "Evaluation of Several Methods for Detecting Sample Mixtures From Different N-dimensional Populations." Paper presented at Amer. Psychol. Assoc. meetings Sept. 9, 1964.

GERARD, R. W. "The Nosology of Schizophrenia: A Cooperative Study." *Behavioral Science*, 1964, 9, 311-333.

GIBSON, W. A. "Three Multivariate Models: Factor Analysis, Latent Structure Analysis, and Latent Profile Analysis." *Psychometrika*, 1959, 24, 229-252.

GIBSON, W. A. "Latent Structure Analysis and Psychiatric Classification." *Trans. Sixth Res. Conf. on Cooperative Chemotherapy Studies in Psychiatry and Research Approaches to Mental Illness*. Washington, D.C.: Veterans Administration, 1961.

GLESER, GOLDINE C. "Logical Constructs Underlying Various Methods of Grouping Persons." *Trans. Sixth Res. Conf. on Cooperative Chemotherapy Studies in Psychiatry and Research Approaches to Mental Illness*. Washington, D.C.: Veterans Administration, 1961.

HOLZINGER, K. J., and HARMAN, H. H. "Factor Analysis." Chicago: Univ. Chicago Press, 1941.

KATZ, MARTIN M. "A Phenomenological Typology of Schizophrenia." In symposium: "Explorations in Typology With Special Reference to Psy-

chotics." Amer. Psychol. Assoc. meetings, Sept. 4, 1964. New York: Human Ecology Fund, Inc.

KATZ, MARTIN M., and COLE, J. O. "A Phenomenological Approach to Classification of Schizophrenic Disorders." *Dis. nerv. Syst.* 1963, 24, 147-154.

LAZARSFELD, P. F. "The Logical and Mathematical Foundations of Latent Structure Analysis." In Stouffer, S. A., et al., "Measurement and Prediction." Princeton: Princeton Univ. Press, 1950.

LORR, M., KLETT, C. J., and MCNAIR, D. M. "Syndromes of Psychosis." New York: McMillan, 1963.

LORR, M. "Some Syndrome-based Psychotic Types." In symposium: "Explorations in Typology with Special Reference to Psychotics." Amer. Psychol. Assoc. meetings, Sept. 4, 1964. New York: Human Ecology Fund, Inc.

MAHALANOBIS, P. C. "On the Generalized Distance in Statistics." *Proc. Nat. Inst. Sci., India*, 1936, 12, 48-58.

MCQUITTY, L. L. "Pattern Analysis Illustrated in Classifying Patients and Normals." *Educ. psychol. Measmt.*, 1954, 14, 598-604.

MCQUITTY, L. L. "Hierarchical Syndrome Analysis." *Educ. psychol. Measmt.*, 1960, 20, 805-816.

MCQUITTY, L. L. "Typal analysis." *Educ. psychol. Measmt.*, 1961, 21, 677-696.

MCQUITTY, L. L. "Rank order typal analysis." *Ed. & psych. Measmt.*, 1963, 23, 55-61.

MCNAIR, D. M., and LORR, M. "Differential Typing of Psychotic Outpatients." *Psychol. Rec.*, 1965, 15, 33-41.

NUNNALLY, JUM. "The Analysis of Profile Data." *Psychol. Bull.*, 1962, 59, 311-319.

OVERALL, JOHN E. "Applications of Profile Classification Procedures in Psychiatric Drug Research." In symposium: "Explorations in Typology With Special Reference to Psychotics." Amer. Psychol. Assoc. meetings, Sept. 4, 1964. New York: Human Ecology Fund, Inc.

OVERALL, JOHN E. "Note on Multivariate Methods for Profile Analysis." *Psychol. Bull.*, 1964, 61, 195-198.

PEARSON, K. P. "Contributions to the Mathematical Theory of Evolution." "I. On the Dissection of Asymmetrical Frequency Curves." *Phil. Trans. Roy. Soc., A*, 1894, 185, 71-90.

RAO, C. R. "Advanced Statistical Methods in Biometric Research." New York: Wiley, 1952.

SAUNDERS, D. R. "Some Varieties of Schizophrenia." In symposium: "Explorations in Typology With Special Reference to Psychotics." Amer. Psychol. Assoc. meetings, Sept. 4, 1964. New York: Human Ecology Fund, Inc.

SAUNDERS, D. R., and SCHUCMAN, HELEN. "Syndrome analysis: An efficient procedure for isolating meaningful subgroups in a population." Paper presented at Psychonomic Society, Sept. 1, 1962.

- THORNDIKE, R. L. "Who belongs in the family?" *Psychometrika*, 1953, 18, 267-276.
- WARD, JOE H., Jr. "Hierarchical Grouping to Optimize an Objective Function." *J. Amer. Statis. Assoc.*, 1963, 58, 236-244.
- WARD, JOE H., Jr., and HOOK, MARION E. "Applications of an Hierarchical Grouping Procedure to a Problem of Grouping Profiles." *Educ. Psychol. Measmt.*, 1963, 23, 69-81.
- ZUBIN, J., and FLEISS, J. L. "Taxonomy in the Mental Disorders—A Historical Perspective." In symposium: "Explorations in Typology With Special Reference to Psychotics." *Amer. Psychol. Assoc. meetings*, Sept. 4, 1964. New York: Human Ecology Fund, Inc.
- ZUBIN, J., FLEISS, J. L., and BURDICK, E. I. "A Method for Fractionating a Population Into Homogeneous Subgroups." Unpublished paper, 1963.

Discrimination Among Groups and Assigning New Individuals

C. Radhakrishna Rao, Sc. D., F.R.S.¹

1. Introduction

The organizers of this conference made my task easy by asking me to speak on a specific problem, viz., that of "assigning new individuals as members of certain well-defined groups." I shall discuss some recent work in this direction and indicate the areas where further research is necessary.

An extended formulation of this problem recognizes the possibility that a given individual may not belong to any of the specified groups but to an unknown group whose existence has not been established earlier. Thus, rules have to be formulated for assigning a new individual to one of a specified set of groups or to none. Such decision rules enable us to bring to light new groups by the chance occurrence of individuals from them.

A generalization of the classical problem is the assignment of an observed individual

as a member of one of specified clusters of groups. We shall first consider the classical problem and then discuss some extensions and generalizations.

The problem of assigning a new individual to one of a finite number of groups to which he may belong is referred to in statistical literature as one of classification or discrimination. It has been recently suggested that *identification* is a more appropriate terminology.

In the statistical approach to the problem, we first characterize each group (of the possible groups) by the distribution of certain measurements on the individuals of that group. The characters must be such that their distributions in the different groups are all different. Then, on the basis of the measurements ascertained on a new individual, a decision rule (a procedure) is provided for deciding on the individual's membership in one of the groups. We shall first consider the situation in which a satisfactory solution is available.

Let \mathbf{X} denote the vector of measurements (random variable) and $f_1(\mathbf{x}), \dots, f_k(\mathbf{x})$ the probability density functions of \mathbf{X} in the k groups. Furthermore, let us consider a new individual to be assigned to one of the groups as randomly observed from a mixed population consisting of individuals of the k groups in known proportions π_1, \dots, π_k . The quantities π_1, \dots, π_k are referred to as prior probabilities and if the value of \mathbf{X} on an observed individual is \mathbf{x} , then by applying Bayes theorem the posterior probabilities of the k groups (given \mathbf{x}) are

$$(1.1) \quad \frac{\pi_1 f_1(\mathbf{x})}{\sum \pi_i f_i(\mathbf{x})}, \dots, \frac{\pi_k f_k(\mathbf{x})}{\sum \pi_i f_i(\mathbf{x})}$$

Knowing these probabilities, the consequences of any decision procedure can be examined.

A general decision rule is to throw a k faced die with probabilities

$$(1.2) \quad \lambda_1(\mathbf{x}), \dots, \lambda_k(\mathbf{x}); \sum \lambda_i(\mathbf{x}) = 1$$

for the k faces, depending on the observed value \mathbf{x} , and decide on the i^{th} group if the

¹ Director, Research and Training School, Indian Statistical Institute, Calcutta, India.

i^{th} face appears. Such a procedure is known as a randomized decision rule. If r_{ij} is the loss resulting in assigning a member of the i^{th} group to the j^{th} group then the expected loss for given \mathbf{x} is

$$(1.3) \quad = \sum_i \sum_j \pi_i f_i(\mathbf{x}) \lambda_j(\mathbf{x}) r_{ij} / \sum_i \pi_i f_i(\mathbf{x}) \\ - [\lambda_1(\mathbf{x}) S_1(\mathbf{x}) + \dots + \lambda_k(\mathbf{x}) S_k(\mathbf{x})]$$

where

$$(1.4) \quad S_i(\mathbf{x}) = - \sum_j r_{ji} \pi_j f_j(\mathbf{x}) / \sum_i \pi_i f_i(\mathbf{x})$$

is called the i^{th} discriminant score. It is clear that the expected loss (1.3) is a minimum when the $\lambda(\mathbf{x})$ corresponding to the highest discriminant score is unity and zero otherwise. Or, when there is no unique highest discriminant score the probabilities $\lambda_i(\mathbf{x})$ corresponding to the highest scores can be chosen arbitrarily while the probabilities for the rest are chosen to be zero.

In a wide variety of problems we may choose $r_{ii} = 0$, $r_{ij} = 1$ for $i \neq j$. In such a case, the expected loss corresponds to the expected proportion of wrong identifications. The i^{th} discriminant score is

$$(1.5) \quad S_i(\mathbf{x}) = [- \sum_j \pi_j f_j(\mathbf{x}) \\ + \pi_i f_i(\mathbf{x})] \div \sum_j \pi_j f_j(\mathbf{x})$$

Since the decision rule involves only a comparison of the S_i values, we may drop the constant term $- \sum_j \pi_j f_j(\mathbf{x})$ in (1.5) and define the discriminant score simply as

$$(1.6) \quad S_i(\mathbf{x}) = \pi_i f_i(\mathbf{x})$$

which is proportional to the posterior probability of the i^{th} group given the observation \mathbf{x} .

A slightly different choice of the loss elements r_{ij} is $r_{ij} = r_i$ for $i \neq j$ and $r_{ii} = 0$, i.e., the loss essentially depends on the group to which an individual belongs but not on the particular wrong group to which he is assigned. In such a case

$$(1.7) \quad S_i(\mathbf{x}) = [- \sum_j r_j \pi_j f_j(\mathbf{x}) \\ + r_i \pi_i f_i(\mathbf{x})] \div \sum_j \pi_j f_j(\mathbf{x})$$

Dropping the constant terms in (1.7) we may define the i^{th} discriminant score simply as

$$(1.8) \quad S(\mathbf{x}) = r_i \pi_i f_i(\mathbf{x})$$

Whatever may be the choice of the discriminant scores like (1.4) or (1.6) or (1.8), the decision rule is as follows:

(a) If there is a unique highest discriminant score among $S_1(\mathbf{x})$, \dots , $S_k(\mathbf{x})$, then assign the individual to that group for which the discriminant score is the highest.

(b) If there is more than one group for which the discriminant scores are equal and the highest, then assign the individual arbitrarily to any one of such groups.

2. Application of the Optimum Rule

There are a number of difficulties in the application of the optimum decision rule in practice.

(i) The quantities needed, viz., the prior probabilities

$$(2.1) \quad \pi_1, \dots, \pi_k$$

and the density functions

$$(2.2) \quad f_1, \dots, f_k$$

for an application of the optimum decision rule may not be known. However, they may be estimable from data suitably collected.

(ii) While the optimum decision rule lays down a strict procedure of coming to a decision in all cases, some caution is necessary in practice. If one discriminant score is considerably large compared to the others, there need not be any mental reservation in arriving at a decision. On the other hand, when the next highest discriminant scores are close to the highest value, it is probably wise to consider the individual as belonging to one of the groups in a subset chosen on the basis of the highest discriminant scores and look for further evidence. What is the exact procedure to be followed in such cases?

(iii) In practice, it may not be possible or desirable to obtain all the measurements on an individual referred to for identification in one stage. A sequential approach with the possibility of arriving at a decision before all the measurements are completed is desirable. Such a procedure may result

in considerable saving (in the long run) of the measurements some of which may be very expensive to obtain.

(iv) What are the considerations for arriving at the best choice and sequence of measurements?

(v) The a priori information that an observed individual belongs to one of the k given groups may be wrong. In fact he may belong to another unknown group and it is, therefore, necessary to develop a theory which takes into account such a possibility. This is important in two ways. Firstly, it may enable us to discover a new group whose existence in the population under consideration has not yet been established. Secondly, it admits the possibility of discovering any contamination taking place in the population (on which we are applying a decision rule) by the injection of individuals from an outside group. An example is the discovery of cholera cases in Japan a couple of years ago. The symptoms might have been mistaken as rare manifestations of one of the ailments ordinarily occurring in Japan if the possibility of a sporadic contamination from an outside source had not been kept in mind.

We shall consider these difficulties one by one and suggest suitable modifications in the procedure laid down in section 1 of the paper.

2a. Estimation of the unknown quantities

Estimation of the density functions

In practice, it may be possible to obtain a sample of individuals from each identified group and estimate the distribution of a set of measurements in each group. Sometimes we may have a sample from a population with individuals from the different groups mixed in unknown proportions, in which case we have to employ special techniques for estimating the individual distributions and the proportions of mixture. The method of estimation in any case depends on the nature of measurements (qualitative, quantitative, or mixed) and the assumptions

regarding the specification of their distribution. In the case of quantitative data and normality assumption we need only estimate the mean values and the dispersion matrix which completely determine the density function. The situation is somewhat complicated with qualitative data. There have been two approaches to the problem, one suggested by Bahadur (1961), and another by Lazarsfeld (1961).

When the estimated density functions are used in the place of the true functions in the decision rules derived in section 1, the expected loss will be somewhat larger than the minimum attainable, depending on the errors of estimation (i.e., on the sizes of the samples providing information on the unknown parameters). Some investigations on the loss incurred due to estimation for various sample sizes have been made by Cochran and Hopkins (1961) in the case of qualitative data and by John (1961) in the case of quantitative data. The extra error is small when the estimates of parameters are based on large samples. In practice, as observations accumulate more precise estimates of parameters will become available and the loss due to errors of estimation thus gets continuously diminished.

Estimation of prior probabilities

The optimum rule depends on the relative frequencies π_1, \dots, π_k of the individuals of the k different groups in the population from which an individual to be identified is drawn (π_1, \dots, π_k may also be considered as the relative frequencies with which individuals from different groups present themselves for identification). At the beginning, we may have only crude estimates of π_i but as the measurements on individuals referred to for identification accumulate, precise estimates will be available leading to improved decision rules.

For estimating π_1, \dots, π_k let us assume that the density functions f_1, \dots, f_k are known. Let $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ be the observations on n individuals referred to for identi-

fication. Then the probability density at the observed values is

$$(2.3) \quad L = \prod_{i=1}^n [\pi_1 f_1(\mathbf{x}_i) + \cdots + \pi_k f_k(\mathbf{x}_i)]$$

We may then apply the method of maximum likelihood to estimate π_1, \dots, π_k . But the computations will be extremely heavy. Therefore, we suggest a simpler method which depends only on the mean values and the covariance matrix.

Let $\pi_{10}, \dots, \pi_{k0}$ be the provisional values of the prior probabilities, $\mathbf{u}_1, \dots, \mathbf{u}_k$, be the mean values in the k groups and $\mathbf{\Lambda}_1, \dots, \mathbf{\Lambda}_k$ be the dispersion matrices in the k groups. Further let $\bar{\mathbf{X}}$ denote the average vector of measurements on n new individuals and π_1, \dots, π_k the prior probabilities to be determined. We observe that

$$(2.4) \quad E(\bar{\mathbf{X}}) = \pi_1 \mathbf{u}_1 + \cdots + \pi_k \mathbf{u}_k$$

$$(2.5) \quad D(\bar{\mathbf{X}}) = \frac{1}{n} (\pi_1^2 \mathbf{\Lambda}_1 + \cdots + \pi_k^2 \mathbf{\Lambda}_k)$$

where E denotes expectation and D , dispersion (variance-covariance) matrix. A crude estimate of $D(\bar{\mathbf{X}})$ is

$$(2.6) \quad \frac{1}{n} (\pi_{10}^2 \mathbf{\Lambda}_1 + \cdots + \pi_{k0}^2 \mathbf{\Lambda}_k)$$

To estimate π_1, \dots, π_k , we may apply the theory of least squares (with the restrictions, $\pi_i \geq 0$ and $\sum \pi_i = 1$) on the observation vector $\bar{\mathbf{X}}$ with its expectation as in (2.4) where \mathbf{u}_i are known and with the (approximate) covariance matrix as in (2.6). We may choose the estimates so computed as a second set of provisional estimates for obtaining a better approximation to the dispersion matrix. The least squares method is repeated with the new dispersion matrix. There is no need to repeat this operation beyond two stages specially when there is a further opportunity of collecting more data and revising the estimates.

In the problem of differential diagnosis of diseases, the frequencies of individuals likely to suffer from different diseases may change over time and they may even exhibit

seasonal variations during the course of a year. Estimates appropriate to a given season and point of time have to be used to obtain the best possible results. The estimation of relative frequencies as functions of time may not be easy. Some research is necessary in this direction.

2b. Choice of characters

Theoretically the addition of each extra character can possibly improve the decision rule but not make it worse. But in practice, when the exact values of the parameters needed for setting up the optimum decision rule are unknown but only estimates are available, the use of some additional characters in the usual way may result in loss of efficiency. Therefore, there is a need for making a suitable choice of characters.

We have to guard against two different types of situations. First, some of the characters may be nearly (apart from minor technical errors) linearly dependent. This causes difficulties in numerical computations specially when dispersion matrices have to be inverted. Such a difficulty can be avoided by omitting some characters (or linear combinations of characters). Second, certain characters (or linear combinations of characters) may not show variation between groups independently of the others. The inclusion of such characters (or linear combinations of characters) will not be useful. There are two ways of examining the problem.

Let B and W be the dispersion matrices between and within groups for the standardized characters, based on samples from the different groups. Further let

$$(2.7) \quad \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p$$

be the eigen values and

$$(2.8) \quad \mathbf{m}_1, \mathbf{m}_2, \dots, \mathbf{m}_p$$

be the corresponding eigen vectors of the "within matrix" W , where p is the number of measurements. The eigen vectors provide a linear transformation of the original measurements to a set uncorrelated variables

$$(2.9) \quad Y_i = \mathbf{m}_i' \mathbf{X}, \quad i = 1, \dots, p$$

where we use the same symbol \mathbf{X} for the standardized original measurements. Now we compute the quantities

$$(2.10) \quad \partial_i = \frac{\mathbf{m}_i' \mathbf{B} \mathbf{m}_i}{\lambda_i}, \quad i = 1, \dots, p$$

which indicate the relative importance of the functions $\mathbf{m}_i' \mathbf{X}$ in discrimination.

A small value of λ indicates that the corresponding linear function $\mathbf{m}' \mathbf{X}$ is nearly a constant and a small value of ∂ indicates that the corresponding linear function is a poor discriminator. The analysis we have undertaken thus enables us to examine the two points raised. We may omit linear functions with small values of ∂ . The decision procedure is set up using the remaining linear combinations as new measurements.

If there is a linear relationship among the characters, which may be obscured by errors in measurements, it would be indicated by a small value of λ as well as the corresponding ∂ . On the other hand there may be a large value of ∂ associated with a small value of λ . This does not imply a linear relationship of the type we have indicated but some caution is necessary in such a case. Occasionally spurious values may arise due to approximations made in the computations specially when λ is small.

If the object is to omit the superfluous measurements but not suitable linear combinations as was possible by determining the eigen values and vectors of W , then a different transformation of the measurements, known as Gram-Schmidt orthogonalization, is necessary. In this method linear functions are constructed starting from the first variable and adding on the other variables in stages:

$$(2.11) \quad \begin{aligned} y_1 &= x_1 \\ y_2 &= a_{21}x_1 + x_2 \\ &\vdots \\ y_i &= a_{i1}x_1 + \dots + a_{i, i-1}x_{i-1} + x_i \\ &\vdots \end{aligned}$$

At the i^{th} stage we compute the variance of y_i between and within groups. Let λ_i

denote the variance within and ∂_i the ratio of between to within. A decision to drop x_i is taken if ∂_i is small. The process of building up the linear functions is continued omitting such x_i . Finally, the decision procedure is set up using the linear functions so constructed.

2c. Sequential decision rules

Let $S_i^{(j)}$ be the discriminant score for the i^{th} group based on the first j measurements. A sequential decision rule is of the following type:

(a) Stop further measurements after the j^{th} if

$$(2.12) \quad \max \{ S_1^{(j)}, \dots, S_k^{(j)} \} \geq S_o$$

and if $S_i^{(j)}$ is the maximum, assign the individual to the i^{th} group.

(b) Take additional measurements if

$$(2.13) \quad \max \{ S_1^{(j)}, \dots, S_k^{(j)} \} < S_o$$

(c) If no decision is reached before the p^{th} measurement, then assign the individual to i^{th} group if

$$(2.14) \quad S_i^{(p)} \geq S_1^{(p)}, \dots, S_k^{(p)}$$

If it were possible to continue taking further measurements till the condition (2.12) is satisfied, then the decision rule has the property that the expected risk is smaller than $(-S_o)$. But if the process is truncated at the p^{th} stage and rule (c) is adopted, the expected risk will be larger than $(-S_o)$. The exact computation of the expected risk would be difficult. But it may be possible to obtain some idea by Montecarlo techniques.

The optimum sequence of measurements depends on the costs of making the different measurements and the discriminatory power of different subsets of the measurements. The sequence for which the total of the cost making the measurements and the loss of wrong assignments is a minimum at each stage has to be preferred.

2d. Detection of "outside contamination"

As mentioned earlier, we should keep open the possibility that a new individual

does not belong to any of the k specified groups. That is, we should incorporate in our decision rule a situation in which we decide that a new individual belongs to an unknown group.

Let us suppose that the measurements have a p -variate normal density in each of the k specified groups and also in the unknown group to which an individual may belong. Then, on the basis of measurements \mathbf{x} on an individual, the likelihood ratio criterion for testing the hypothesis that the individual is a member of one of the k specified groups is

$$(2.15) \quad f(\mathbf{x}) = \pi_1 f_1(\mathbf{x}) + \cdots + \pi_k f_k(\mathbf{x})$$

If $f(\mathbf{x}) < c$ (where c is chosen such that the level of significance has a specified value), we reject the hypothesis and decide that the individual belongs to an outside group. If $f(\mathbf{x}) \geq c$, then we apply the optimum decision rule for deciding the individual's membership in one of the k specified groups. The determination of c for a given level of significance does not seem to be easy. It would be worth examining whether c can be determined approximately in a simple way.

The problem may also be examined in an alternative way. Let us suppose that the mean vector of the measurements is \mathbf{u}_i in the i^{th} group and that the dispersion matrix is the same in all the groups and is equal to Λ . Let us consider a test of the hypothesis that the new individual belongs to a group with its mean vector as

$$(2.16) \quad \lambda_1 \mathbf{u}_1 + \cdots + \lambda_k \mathbf{u}_k$$

where $\lambda_1, \dots, \lambda_k$ are unknown but subject to the condition $\sum \lambda_i = 1$. The hypothesis does not necessarily specify that the new individual belongs to one of the k groups. It keeps open the possibility that he may belong to an outside group which is related to the specified groups in a special way, as indicated by the equation (2.16) connecting the mean values. To test the hypothesis (2.16) we consider the test criterion

$$(2.17) \quad \chi^2 = \min (\mathbf{x} - \lambda_1 \mathbf{u}_1 - \cdots - \lambda_k \mathbf{u}_k)' \Lambda^{-1} (\mathbf{x} - \lambda_1 \mathbf{u}_1 - \cdots - \lambda_k \mathbf{u}_k)$$

where minimization is with respect to $\lambda_1, \dots, \lambda_k$ subject to the condition $\sum \lambda_i = 1$. The statistic (2.17) has a chi-square distribution on $(p - k + 1)$ degrees of freedom, when the measurements have a p -variate normal distribution.

If the χ^2 is significant at a chosen level of significance, then we decide that the individual belongs to an outside group. Then we compute

$$(2.18) \quad \chi_i^2 = (\mathbf{x} - \mathbf{u}_i)' \Lambda^{-1} (\mathbf{x} - \mathbf{u}_i) - \chi^2, \quad i = 1, \dots, k$$

where χ^2 is as in (2.17). The statistic (2.18) for given i is distributed as χ^2 on $(k-1)$ degrees of freedom on the hypothesis that the individual belongs to the i^{th} group. If

$$(2.19) \quad \chi_i^2 > c, \quad i = 1, \dots, k$$

then, again, we decide that the individual belongs to an outside group which is related to the given groups in the manner indicated in (2.16). The value of c is determined such that the level of significance has a given value (for the null hypothesis that the individual belongs to one of the k given groups). If such a null hypothesis is not rejected, we decide to assign the individual to the i^{th} group if

$$(2.20) \quad \chi_i^2 = \min \{ \chi_1^2, \dots, \chi_k^2 \}$$

For further results, details of some of the methods discussed in the present paper and applications the reader is referred to papers by Fisher (1936, 1939) and Rao (1948, 1952, 1954a, 1954b, 1960, 1962).

The procedure suggested in (2.19) and (2.20) does not involve the priori probabilities. If the priori probabilities are known, then we proceed as follows. Let us represent by \mathbf{T} , the vector random variable $(f_2/f_1, \dots, f_k/f_1)$. It is known (Rao, 1962) that \mathbf{T} is sufficient for the set of populations with mean values of the form

$$(2.21) \quad \lambda_1 \mathbf{u}_1 + \cdots + \lambda_k \mathbf{u}_k$$

Let $P_1(\mathbf{t}), \dots, P_k(\mathbf{t})$ be the probability densities of \mathbf{T} according to the k specified groups and π_1, \dots, π_k the corresponding

prior probabilities. Then instead of the statistic (2.19) we use the test criterion

$$(2.22) \quad \pi_1 P_1(\mathbf{t}) + \cdots + \pi_k P_k(\mathbf{t}) < a$$

where a is chosen such that the level of significance has a given value. If the observed \mathbf{t} satisfies (2.22), then again we decide that the observed individual belongs to an outside group. Otherwise we use the optimum rule of section 1 in assigning the individual to one of the specified groups.

3. Discriminant Function Between Composite Hypotheses and Related Problems

The discriminant function, as introduced by the late Sir Ronald Fisher, for deciding between two simple hypotheses (alternative populations) on the basis of observed data is the logarithm of the likelihood ratio of two simple hypotheses given the observations (Welch, 1939). It is known that the discriminant function so obtained provides a sufficient reduction of data for drawing inferences on the two alternative hypotheses (Smith, 1949). In some situations a discriminant function derived from two simple hypotheses may provide a sufficient reduction of data for drawing inferences on a wider set of alternatives (Rao, 1962). The importance of the latter result in practical applications of the discriminant function was demonstrated in earlier publications (Rao, 1961, 1962, and 1965).

The question naturally arises as to what is a suitable discriminant function when the alternative hypotheses are not simple but composite. Such a problem was faced by Burnaby (1965) when he wanted to identify an individual as belonging to one of two sets of populations. Each set consisted of several populations (mixed in unknown proportions) of organisms of one kind representing different (unknown) stages of growth. The object was to decide as to which of two kinds a given organism belongs when nothing is known about its stage of growth. The problem becomes different when some indicator of the organism's stage of growth is available (see De-

lany, 1964; Delany and Healy, 1964). The present section is devoted to a general discussion of the former type of problems.

3a. Discrimination of composite hypotheses: general methods

Let \mathbf{X} denote a random variable and $P(\cdot | \theta)$ the density function depending on a parameter θ belonging to a set \mathbb{H} . Let H_1 be the hypothesis that $\theta \in \mathbb{H}_1$, and H_2 be the hypothesis that $\theta \in \mathbb{H}_2$, where \mathbb{H}_1 and \mathbb{H}_2 are exclusive subsets of \mathbb{H} . The problem we consider is that of choosing between the composite hypotheses H_1 and H_2 on the basis of an observed value of X . Let us discuss a few possible approaches to the problem.

Decision theory set up

Let $F(\theta)$ be the prior distribution function of θ , $l_1(\theta)$ be the loss in choosing H_1 when $\theta \in \mathbb{H}_2$ and $l_2(\theta)$ be the loss in choosing H_2 when $\theta \in \mathbb{H}_1$. Given an observation $X = x$, the conditional expected loss in deciding on H_1 is proportional to

$$(3.1) \quad r_1(x) = \int_{\mathbb{H}_2} l_1(\theta) P(x | \theta) dF(\theta)$$

and that in deciding on H_2 is proportional to

$$(3.2) \quad r_2(x) = \int_{\mathbb{H}_1} l_2(\theta) P(x | \theta) dF(\theta)$$

Then it is easily seen that the expected loss is a minimum if the following decision rule is chosen. "Accept H_2 if $r_1(x) > r_2(x)$, H_1 if $r_1(x) < r_2(x)$ and any one at random if $r_1(x) = r_2(x)$." The discriminant function is then $r_1(x) - r_2(x)$. The solution is not practically acceptable if $l_i(\theta)$ and $F(\theta)$ are not known. So we try other methods.

Solution based on similar divisions

Let R_1 and R_2 be two exclusive regions covering the entire sample space. The regions R_1 , R_2 are said to provide a similar division of the space if there exist constants e_1 , e_2 such that

$$(3.3) \quad \int_{R_2} P(x | \theta) dx = e_1 \text{ for each } \theta \in \mathbb{H}_1$$

and

$$(3.4) \quad \int_{R_1} P(x | \theta) dx = e_2 \text{ for each } \theta \in \Theta_2$$

Let us decide to choose H_1 if $x \in R_1$ and H_2 if $x \in R_2$. In such a case the errors committed are e_1 and e_2 . For determining an optimum decision rule, we consider all similar divisions and choose the one for which the magnitudes of errors are the smallest subject to a given ratio of errors, or for which a given linear compound of errors is a minimum.

There are two ways of arriving at such a solution. Let T be a sufficient statistic (function of X) for θ restricted to Θ_1 , and let the same statistic be sufficient also for θ restricted to Θ_2 . Using the well known factorization theorem (see p. 49, Lehmann, 1959) we may write

$$(3.5) \quad \begin{aligned} P(x | \theta) &= P(t | \theta) P_1(x | t), & \theta \in \Theta_1 \\ &= P(t | \theta) P_2(x | t), & \theta \in \Theta_2 \end{aligned}$$

where the functions $P_1(x | t)$ and $P_2(x | t)$ are independent of θ and may be interpreted as conditional densities of the observations given $T = t$.

If we choose two values $\theta_1 \in \Theta_1$ and $\theta_2 \in \Theta_2$, then the discriminant function for distinguishing between θ_1 and θ_2 is $\log P(x | \theta_1) / P(x | \theta_2)$. Using the factorizations (3.5) we have

$$(3.6) \quad \frac{P(x | \theta_1)}{P(x | \theta_2)} = \frac{P(t, \theta_1)}{P(t, \theta_2)} \cdot \frac{P_1(x | t)}{P_2(x | t)}$$

Taking logarithms

$$(3.7) \quad \log \frac{P(x | \theta_1)}{P(x | \theta_2)} = \log \frac{P(t, \theta_1)}{P(t, \theta_2)} + \log \frac{P_1(x | t)}{P_2(x | t)}$$

which provides a decomposition of the discriminant function for the simple hypotheses θ_1, θ_2 as the sum of two discriminant functions, one based t alone and another on the conditional distributions given t .

It is easy to see that the second component of (3.7) has the same distribution for all θ belonging to any particular set

Θ_1 or Θ_2 , and so it does not discriminate between parameter values within a given set. When the conditional densities $P_1(x | t)$ and $P_2(x | t)$ are different, we have discrimination between parameter values belonging to the different sets (or between the hypotheses H_1 and H_2) by using the discriminant function $\log [P_1(x | t) / P_2(x | t)]$.

Note that the success of the method depends on the conditional density functions $P_1(x | t)$ and $P_2(x | t)$ being different. If T happens to be sufficient for θ over the entire range $\Theta_1 \cup \Theta_2$, then $P_1(x | t)$ and $P_2(x | t)$ are the same and the equation (3.7) merely shows that the discriminant function between two simple hypotheses is an explicit function of the sufficient statistic.

Solution based on ancillary statistics

Another method is to consider a statistic S (function of X) such that its probability density,

$$(3.8) \quad \begin{aligned} P(s | \theta) &= P_1(s) \text{ independent of } \theta \in \Theta_1 \\ &= P_2(s) \text{ independent of } \theta \in \Theta_2 \end{aligned}$$

or, in other words, S is an ancillary statistic for $\theta \in \Theta_1$ and also for $\theta \in \Theta_2$. When $P_1(s)$ and $P_2(s)$ are different, the discriminant function for choosing between H_1 and H_2 is provided by the likelihood ratio $P_1(s) / P_2(s)$.

Method of maximum likelihood ratio

A discriminant function which may have wide applicability is the ratio

$$(3.9) \quad \sup_{\theta \in \Theta_1} P(x | \theta) \div \sup_{\theta \in \Theta_2} P(x | \theta)$$

(see p. 15, Lehmann, 1959).

It will be difficult to give a general discussion of the applicability or of the relative performances of the various suggested procedures. We shall, therefore, consider some special cases which have important applications. In these special cases the various approaches lead to the same discriminant function.

3b. Discrimination of composite hypotheses: special cases

Let us consider the special case where \mathbf{X} has a p -variate normal distribution.

Problem 1.—Let H_1 and H_2 be defined as follows, where E and D stand for expectation and dispersion operators, respectively.

$$(3.10) \quad \begin{aligned} H_1: E(\mathbf{X}) &= \mathbf{a}_1 + \mathbf{B}'\theta_1, D(\mathbf{X}) = \Lambda \\ H_2: E(\mathbf{X}) &= \mathbf{a}_2 + \mathbf{B}'\theta_2, D(\mathbf{X}) = \Lambda \end{aligned}$$

where \mathbf{a}_1 and \mathbf{a}_2 are p -vectors, θ_1, θ_2 are k -vectors and \mathbf{B}' is $p \times k$ matrix of rank k . The values of $\mathbf{a}_1, \mathbf{a}_2$ and \mathbf{B}' are fixed but those of θ_1, θ_2 are arbitrary. The H_1 and H_2 are composite hypotheses.

For example, each composite hypothesis may consist of populations representing various stages of growth of an organism. The mean of any character X_i (the i^{th} component of \mathbf{X}) for organism with age t may be written $E(X_i) = \alpha_i + \beta_i t$, where β_i is the regression coefficient with time. The regression coefficients β_i are taken to be the same for two sets of populations but α_i may be different. The problem is to identify an organism as belonging to one of two sets of populations when the age of the organism is not known.

Considering the general case of (3.10) it is easy to verify that the statistic $\mathbf{B}\Lambda^{-1}\mathbf{X}$ is sufficient for θ_1 and also for θ_2 . Let $\delta = (\mathbf{a}_1 - \mathbf{a}_2) + \mathbf{B}'\phi$, where $\phi' = \theta_1 - \theta_2$. Then

$$(3.11) \quad \begin{aligned} E(\mathbf{B}\Lambda^{-1}\mathbf{X} | H_1) \\ - E(\mathbf{B}\Lambda^{-1}\mathbf{X} | H_2) = \mathbf{B}\Lambda^{-1}\delta \end{aligned}$$

$$(3.12) \quad D(\mathbf{B}\Lambda^{-1}\mathbf{X} | H_1 \text{ or } H_2) = \mathbf{B}\Lambda^{-1}\mathbf{B}'$$

The discriminant function based on $\mathbf{B}\Lambda^{-1}\mathbf{X}$ alone is, therefore,

$$(3.13) \quad (\mathbf{B}\Lambda^{-1}\delta)'(\mathbf{B}\Lambda^{-1}\mathbf{B}')^{-1}\mathbf{B}\Lambda^{-1}\mathbf{X}$$

The discriminant function based on the entire observation \mathbf{X} is $\delta'\Lambda^{-1}\mathbf{X}$. Hence applying the result (3.7), the discriminant function based on the conditional distributions of \mathbf{X} given $\mathbf{B}\Lambda^{-1}\mathbf{X}$ is the difference

$$(3.14) \quad \delta'\Lambda^{-1}\mathbf{X} - (\mathbf{B}\Lambda^{-1}\delta)'(\mathbf{B}\Lambda^{-1}\mathbf{B}')^{-1}\mathbf{B}\Lambda^{-1}\mathbf{X}$$

Now, writing $\delta = (\mathbf{a}_1 - \mathbf{a}_2) + \mathbf{B}'\phi$, the expression (3.14) reduces to

$$(3.15) \quad \begin{aligned} &(\mathbf{a}_1 - \mathbf{a}_2)'\Lambda^{-1}\mathbf{X} - (\mathbf{a}_1 - \mathbf{a}_2)'\Lambda^{-1}\mathbf{B}' \\ &\quad (\mathbf{B}\Lambda^{-1}\mathbf{B}')^{-1}\mathbf{B}\Lambda^{-1}\mathbf{X} \\ &= (\mathbf{a}_1 - \mathbf{a}_2)'[\Lambda^{-1} - \Lambda^{-1}\mathbf{B}'(\mathbf{B}\Lambda^{-1} \\ &\quad \mathbf{B}')^{-1}\mathbf{B}\Lambda^{-1}]\mathbf{X} \end{aligned}$$

which depends only on $(\mathbf{a}_1 - \mathbf{a}_2)$ and is independent of ϕ as is to be expected.

To apply the method of ancillary statistics, let us consider the statistic $\mathbf{C}\mathbf{X}$ where \mathbf{C} is $(k - p) \times p$ matrix of rank $(p - k)$ such that $\mathbf{B}\mathbf{C}' = \mathbf{O}$. Then

$$(3.16) \quad \begin{aligned} E(\mathbf{C}\mathbf{X} | H_1) &= \mathbf{C}\mathbf{a}_1, D(\mathbf{C}\mathbf{X} | H_1) = \mathbf{C}\Lambda\mathbf{C}' \\ E(\mathbf{C}\mathbf{X} | H_2) &= \mathbf{C}\mathbf{a}_2, D(\mathbf{C}\mathbf{X} | H_2) = \mathbf{C}\Lambda\mathbf{C}' \end{aligned}$$

under the hypotheses H_1 and H_2 respectively. Thus $\mathbf{C}\mathbf{X}$ is ancillary under the alternatives in H_1 and also in H_2 . The discriminant function based on $\mathbf{C}\mathbf{X}$ is

$$(3.17) \quad (\mathbf{C}\mathbf{a}_1 - \mathbf{C}\mathbf{a}_2)'(\mathbf{C}\Lambda\mathbf{C}')^{-1}\mathbf{C}\mathbf{X}$$

It may be seen that (3.15) and (3.17) are the same.

It is easily shown that the method of maximum likelihood ratio as defined in (3.9) also yields the same discriminant function.

Problem 2.—In problem 1, the dispersion matrices under the two hypotheses were the same. Let us now consider the alternative composite hypotheses

$$(3.18) \quad \begin{aligned} H_1: E(\mathbf{X}) &= \mathbf{a}_1 + \mathbf{B}'\theta_1, D(\mathbf{X}) = \Lambda_1 \\ H_2: E(\mathbf{X}) &= \mathbf{a}_2 + \mathbf{B}'\theta_2, D(\mathbf{X}) = \Lambda_2 \end{aligned}$$

where θ_1, θ_2 are arbitrary as in problem 1.

It is easily seen that $\mathbf{B}\Lambda_1^{-1}\mathbf{X}$ is sufficient for θ_1 while $\mathbf{B}\Lambda_2^{-1}\mathbf{X}$ is sufficient for θ_2 . Since the two sufficient statistics are not the same, the method of conditional distributions cannot be applied, unless one considers the statistic $(\mathbf{B}\Lambda_1^{-1}\mathbf{X}, \mathbf{B}\Lambda_2^{-1}\mathbf{X})$ as jointly sufficient for θ_1 and for θ_2 . But such a statistic is too wide.

But the method of ancillary statistics is applicable since the statistic $\mathbf{C}\mathbf{X}$ where \mathbf{C} is as defined in (3.16), is ancillary under

both the hypotheses. The distributions under H_1 and H_2 are specified by

$$(3.19) \quad \begin{aligned} E(\mathbf{CX} | H_1) &= \mathbf{Ca}_1, D(\mathbf{CX} | H_1) = \mathbf{CA}_1\mathbf{C}' \\ E(\mathbf{CX} | H_2) &= \mathbf{Ca}_2, D(\mathbf{CX} | H_2) = \mathbf{CA}_2\mathbf{C}' \end{aligned}$$

Taking the logarithm of the likelihood ratio we have the discriminant function, $Q(\mathbf{X})$ equal to

$$(3.20) \quad \begin{aligned} &\mathbf{X}'\mathbf{C}'[(\mathbf{CA}_1\mathbf{C}')^{-1} - (\mathbf{CA}_2\mathbf{C}')^{-1}]\mathbf{CX} \\ &- 2[\mathbf{a}_1'\mathbf{C}'(\mathbf{CA}_1\mathbf{C}')^{-1} - \mathbf{a}_2'\mathbf{C}'(\mathbf{CA}_2\mathbf{C}')^{-1}]\mathbf{CX} \end{aligned}$$

which is quadratic in \mathbf{X} . Using the identity

$$(3.21) \quad \begin{aligned} &\mathbf{C}'(\mathbf{CA}_i\mathbf{C}')^{-1}\mathbf{C} = \mathbf{A}_i^{-1} \\ &- \mathbf{A}_i^{-1}\mathbf{B}'(\mathbf{BA}_i^{-1}\mathbf{B}')^{-1}\mathbf{BA}_i^{-1}, \quad i = 1, 2 \end{aligned}$$

we can write (3.20) in terms of \mathbf{B} only. It may be verified that the method of maximum likelihood ratio also provides the same quadratic discriminant function.

3c. Properties of the linear and quadratic discriminant functions

We shall describe the properties of the discriminant functions obtained in section 3b in a series of lemmas.

Lemma 1. Let

$$(3.22) \quad \begin{aligned} E(\mathbf{X} | H_1) &= \mathbf{u}_1 = \mathbf{a}_1 + \mathbf{B}'\mathbf{\theta}_1, & D(\mathbf{X} | H_1) &= \mathbf{\Lambda} \\ E(\mathbf{X} | H_2) &= \mathbf{u}_2 = \mathbf{a}_2 + \mathbf{B}'\mathbf{\theta}_2, & D(\mathbf{X} | H_2) &= \mathbf{\Lambda} \end{aligned}$$

where $\mathbf{a}_1, \mathbf{a}_2, \mathbf{B}, \mathbf{\theta}_1, \mathbf{\theta}_2$, are as defined in (3.10). Further let \mathbf{L} be a p -vector such that $\mathbf{LB} = \mathbf{O}$, and $\mathbf{\delta} = \mathbf{u}_1 - \mathbf{u}_2$. Then

$$(3.23) \quad \sup_{\mathbf{L}} \frac{(\mathbf{L}'\mathbf{\delta})^2}{\mathbf{L}'\mathbf{\Lambda}\mathbf{L}}, \text{ subject to } \mathbf{BL} = \mathbf{O}$$

is attained at

$$(3.24) \quad \mathbf{L}^* = (\mathbf{\Lambda}^{-1} - \mathbf{\Lambda}^{-1}\mathbf{B}'(\mathbf{BA}^{-1}\mathbf{B}')^{-1}\mathbf{BA}^{-1})\mathbf{\delta}$$

The result follows from a general lemma proved in Rao (1964). The function $(\mathbf{L}^*)'\mathbf{X}$ may be called a restricted linear discriminant function whether \mathbf{X} has normal distribution or not, drawing an analogy with the linear discriminant function of Fisher obtained by maximizing the ratio $(\mathbf{L}'\mathbf{\delta})^2/\mathbf{L}'\mathbf{\Lambda}\mathbf{L}$ without any restriction on \mathbf{L} .

Lemma 2. Let $\mathbf{\delta} = \mathbf{a}_1 - \mathbf{a}_2 + \mathbf{B}'\mathbf{\phi}$ as before and $\mathbf{\Lambda} = \mathbf{\Lambda}_o + \mathbf{B}'\mathbf{DB}$ where \mathbf{D} is arbitrary. Then \mathbf{L}^* as determined in lemma 1 is independent of $\mathbf{\phi}$ and \mathbf{D} .

According to lemma 1, \mathbf{L}^* is obtained by considering the class of vectors \mathbf{L} such that $\mathbf{BL} = \mathbf{O}$ and maximizing $(\mathbf{L}'\mathbf{\delta})^2/\mathbf{L}'\mathbf{\Lambda}\mathbf{L}$. Now for such \mathbf{L}

$$(3.25) \quad \begin{aligned} \mathbf{L}'\mathbf{\delta} &= \mathbf{L}'(\mathbf{a}_1 - \mathbf{a}_2) \\ \mathbf{L}'\mathbf{\Lambda}\mathbf{L} &= \mathbf{L}'(\mathbf{\Lambda}_o + \mathbf{B}'\mathbf{DB})\mathbf{L} = \mathbf{L}'\mathbf{\Lambda}_o\mathbf{L} \end{aligned}$$

Therefore the problem is the same as that of maximizing $[\mathbf{L}'(\mathbf{a}_1 - \mathbf{a}_2)]^2/\mathbf{L}'\mathbf{\Lambda}_o\mathbf{L}$ subject to the condition $\mathbf{BL} = \mathbf{O}$. Then the solution is

$$(3.26) \quad \mathbf{L}^* = (\mathbf{\Lambda}_o^{-1} - \mathbf{\Lambda}_o^{-1}\mathbf{B}'(\mathbf{BA}_o^{-1}\mathbf{B}')^{-1}\mathbf{BA}_o^{-1})(\mathbf{a}_1 - \mathbf{a}_2)$$

which is obviously independent of $\mathbf{\phi}$ and \mathbf{D} .

Lemma 2 provides an extremely important result for practical applications. It enables us to construct the discriminant function knowing only the means and dispersion matrices of arbitrary mixtures of populations defined by each composite hypothesis. Thus if $\mathbf{\theta}_1$ has the a priori mean $\bar{\mathbf{\theta}}_1$ and the dispersion matrix \mathbf{D}_1 and $\mathbf{\theta}_2$ has the a priori mean $\bar{\mathbf{\theta}}_2$ and the dispersion matrix \mathbf{D}_2 the mixture of populations under H_1 has the mean and dispersion matrix equal to

$$(3.27) \quad \begin{aligned} \mathbf{u}_1 &= \mathbf{a}_1 + \mathbf{B}'\bar{\mathbf{\theta}}_1, & D(\mathbf{X}) &= \mathbf{\Lambda}_o + \mathbf{B}'\mathbf{D}_1\mathbf{B} \end{aligned}$$

where $\mathbf{\Lambda}_o$ is the dispersion matrix of \mathbf{X} for given $\mathbf{\theta}_1$. Similarly the mixture of populations under H_2 has the mean and dispersion matrix equal to

$$(3.28) \quad \begin{aligned} \mathbf{u}_2 &= \mathbf{a}_2 + \mathbf{B}'\bar{\mathbf{\theta}}_2, & D(\mathbf{X}) &= \mathbf{\Lambda}_o + \mathbf{B}'\mathbf{D}_2\mathbf{B} \end{aligned}$$

Let us consider

$$(3.29) \quad \begin{aligned} \mathbf{\delta} &= \mathbf{u}_1 - \mathbf{u}_2 = \mathbf{a}_1 - \mathbf{a}_2 \\ &+ \mathbf{B}'(\bar{\mathbf{\theta}}_1 - \bar{\mathbf{\theta}}_2) \end{aligned}$$

and

$$(3.30) \quad \begin{aligned} \mathbf{\Lambda} &= w_1 D(\mathbf{X} | H_1) + w_2 D(\mathbf{X} | H_2) \\ &= \mathbf{\Lambda}_o + \mathbf{B}'(w_1\mathbf{D}_1 + w_2\mathbf{D}_2)\mathbf{B} \end{aligned}$$

where w_1 and w_2 are arbitrary weights such that $w_1 + w_2 = 1$. The substitution of $\mathbf{\delta}$

and Λ in the formula (3.24) for L^* gives us a result which is independent of $\bar{\theta}_1, \bar{\theta}_2, D_1, D_2$ and w_1, w_2 . The importance of the result arises due to the fact that in the type of practical situations we are considering we are likely to have only estimates of means and dispersion matrices of (unknown) mixtures of populations in the two groups which we are trying to discriminate.

Lemma 3. Let $\mathbf{u}_1 = \mathbf{a}_1 + \mathbf{B}'\theta_1, \Lambda_1 = \Lambda_{01} + \mathbf{B}'D_1\mathbf{B}$ and $\mathbf{u}_2 = \mathbf{a}_2 + \mathbf{B}'\theta_2, \Lambda_2 = \Lambda_{02} + \mathbf{B}'D_2\mathbf{B}$. Then the quadratic discriminant function, $Q(\mathbf{x})$ as determined in (3.20), is independent of θ_1, θ_2, D_1 and D_2 .

The result is verified by substituting the expressions for $\mathbf{u}_1, \mathbf{u}_2, \Lambda_1, \Lambda_2$ in the formula (3.20). Thus the quadratic discriminant function shares the same property as the linear discriminant function and can be constructed knowing only the means and dispersion matrices of arbitrary mixtures of populations.

3d. Discrimination of several composite hypotheses: canonical analysis

Let us consider several composite hypotheses H_1, \dots, H_k such that

$$(3.31) \quad \begin{aligned} E(\mathbf{X} | H_i) &= \mathbf{a}_i + \mathbf{B}'\theta_i \\ D(\mathbf{X} | H_i) &= \Lambda_i \end{aligned} \quad \begin{aligned} &(\theta_i \text{ arbitrary}) \\ &i = 1, \dots, k \end{aligned}$$

The general theory of section 3a for determining decision rules independent of θ_i applies. In terms of $\mathbf{Y} = \mathbf{C}\mathbf{X}$, the common ancillary statistic for each composite hypothesis, the problem reduces to the discrimination of k simple hypotheses such that

$$(3.32) \quad \begin{aligned} E(\mathbf{Y} | H_i) &= \mathbf{C}\mathbf{a}_i \\ D(\mathbf{Y} | H_i) &= \mathbf{C}\Lambda_i\mathbf{C}' \end{aligned}$$

The solution in such a case follows on standard lines as discussed by Lindley (1953), Rao (1952, 1965), Wald (1950) etc. Observe that in the expression

$$(3.33) \quad \begin{aligned} \log L(H_i | \mathbf{Y}) &= (\mathbf{C}\mathbf{X} - \mathbf{C}\mathbf{a}_i)'(\mathbf{C}\Lambda_i\mathbf{C}')^{-1}(\mathbf{C}\mathbf{X} - \mathbf{C}\mathbf{a}_i) \\ &= (\mathbf{X} - \mathbf{a}_i)' \mathbf{C}'(\mathbf{C}\Lambda_i\mathbf{C}')^{-1} \mathbf{C}(\mathbf{X} - \mathbf{a}_i) \\ &= (\mathbf{X} - \mathbf{a}_i)' [\Lambda_i^{-1} \\ &\quad - \Lambda_i^{-1} \mathbf{B}'(\mathbf{B}\Lambda_i^{-1} \mathbf{B}')^{-1} \mathbf{B}\Lambda_i^{-1}] (\mathbf{X} - \mathbf{a}_i) \end{aligned}$$

we can substitute $\mathbf{a}_i + \mathbf{B}\theta_i$, (where θ_i is arbitrary) for \mathbf{a}_i , and $\Lambda_i + \mathbf{B}'D_i\mathbf{B}$ (where D_i is arbitrary) for Λ_i without altering its value; a result which is of great practical importance.

4. Classification of Groups

It is of some interest to study the inter-relationships among the groups themselves. For this purpose we have to develop the concept of distance between any two groups. A distance function can be defined in many ways. The most satisfactory function in the case of quantitative data is Mahalanobis distance which is related to the frequency of wrong assignments between two groups. In the case of qualitative data, Hellinger's distance can be used.

The first step in the classification of groups is the computation of distances for all possible pairs of groups. The second step consists in choosing pairs of groups which are judged to be close (by the smallness of distance) and building up clusters around them by adjoining other groups. The process of clustering can be continued till the average distance within a cluster does not exceed a suitably chosen value. At the third stage distances between clusters (average distance between members of one cluster with those of another) are computed. The knowledge of the clusters and intra- and inter-cluster distances enable us to examine the inter-relationships among the groups. For details the reader is referred to Rao (1952, ch. 9), Rao (1954a), Rao and Slater (1949), Sokhal and Sneath (1963).

REFERENCES

- (1) BAHADUR, R. (1961): "A Representation of the Joint Distribution of Responses to n Dichotomous Items." Studies in Item Analysis and Prediction. Ed. H. Solomon, pp. 158-169.
- (2) BURNABY, T. P. (1965): "Growth Invariant Discriminant Functions and Generalized Distances." Biometrics (in press).
- (3) COCHRAN, W. G. and HOPKINS, C. E. (1961): "Some Classification Problems With Multivariate

Qualitative Data." *Biometrics* 17, 10-33.

(4) DELANY, M. J. (1964): "Variations in the Long Tailed Field-Mouse (*Apodemus sylvaticus* (L.)) in North-West Scotland. I. Comparison of Individual Characters." *Proc. Roy. Soc. B*, 161, 191-199.

(5) DELANY, M. J., and HEALY, M. J. R. (1964): "Variations in the Long Tailed Mouse, etc. II. Simultaneous Examination of all Characters." *Proc. Roy. Soc. B*, 161, 200-207.

(6) FISHER, R. A. (1936): "The Use of Multiple Measurements in Taxonomic Problems." *Ann. Eugen.* 7, 179-188.

(7) FISHER, R. A. (1939): "The Sampling Distribution of Some Statistics Obtained From Non-Linear Equations." *Ann. Eugen.* 9, 238-249.

(8) JOHN, S. (1961): "Errors in Discrimination." *Ann. Math. Statist.* 32, 1125-1144.

(9) LAZARSFELD, P. F. (1961): "The Algebra of Dichotomous Systems." *Studies in Item Analysis and Prediction*. Ed. H. Solomon, pp. 111-158.

(10) LEHMANN, E. L. (1959): "Testing Statistical Hypotheses." John Wiley, New York.

(11) LINDLEY, D. V. (1953): "Statistical Inference," *J. Roy. Statist. Soc. B*, 11, 30-76.

(12) RAO, C. R. (1948): "The Utilization of Multiple Measurements in Problems of Biological Classification." *J. Roy. Statist. Soc.* 103, 159-203.

(13) RAO, C. R. (1952): "Advanced Statistical Methods in Biometric Research." John Wiley, New York.

(14) RAO, C. R. (1954a): "On the Use and Interpretation of Distance Functions in Statistics." *Bull. Inst. Inter. Statist.* 34, 90-97.

(15) RAO, C. R. (1954b): "A General Theory of Discrimination When the Information on Alternative Hypotheses Is Based on Samples." *Ann. Math. Statist.* 25, 651-670.

(16) RAO, C. R. (1960): "Multivariate Analysis. An Indispensable Aid in Applied Research." *Sankhya* 22, 317-338.

(17) RAO, C. R. (1961): "Some Observations on Multivariate Statistical Methods in Anthropological Research." *Bull. Int. Stat. Inst.*, 38, 99-109.

(18) RAO, C. R. (1962): "Use of Discriminant and Allied Functions in Multivariate Analysis." *Sankhya* 22, 317-338.

(19) RAO, C. R. (1964): "The Use and Interpretation of Principal Component Analysis in Applied Research." *Sankhya* 26, 329-358.

(20) RAO, C. R. (1965): "Linear Statistical Inference and Its Applications." John Wiley, New York.

(21) RAO, C. R., and PATRICK SLATER (1949): "Multivariate Analysis Applied to Differences Between Neurotic Groups." *British J. Psychology* (Stat. Soc.) 2, 17-29.

(22) SMITH, C. A. B. (1947): "Some Examples of Discrimination." *Ann. Eugenics* (London). 18, 272-282.

(23) SOKHAL, R. R. and SNEATH, P. A. (1963): "Principles of Numerical Taxonomy." W. H. Freeman & Co., San Francisco and London.

(24) WALD, A. (1950): "Statistical Decision Functions." John Wiley, New York.

(25) WELCH, B. L. (1939): "Note on Discriminant Functions." *Biometrika* 31, 213.

Discussion

SAMUEL W. GREENHOUSE, Ph.D., *Chief, Theoretical Statistics and Mathematics Section, Office of Biometry, National Institute of Mental Health, Chairman*

(Papers: "On the Meaning of Discrimination, Classification, Mixture, and Clustering in Statistics," Samuel W. Greenhouse, Ph.D., Chief, Theoretical Statistics and Mathematics Section, Office of Biometry, National Institute of Mental Health. "Quantifying Similarity Between People," Goldine C. Gleser, Ph.D., Professor of Psychology, Department of Psychiatry, University of Cincinnati College of Medicine. "Multidimensional Representation of Similarity Structures," Warren S. Torgerson, Ph.D., Department of Psychology, Johns Hopkins University. "A Survey of Some Empirical Clustering Procedures," Samuel B. Lyerly, Ph.D., Psychological Assessment Associates, and Bureau of Social Science Research, Inc., Washington, D.C. "Discrimination Among Groups and Assigning New Individuals," C. Radhakrishna Rao, Sc.D., F.R.S., Director, Research and Training School, Indian Statistical Institute, Calcutta)

OPEN DISCUSSION

(Dr. Greenhouse served as moderator of the discussion.)

Dr. FLEISS. Dr. Greenhouse, you suggested that psychologists or psychometricians can go ahead and apply the available clustering procedures because there seem to be no better procedures available, but you also say, "Don't take the results too seriously." Now, this makes such exercises nothing but exercises, and empty exercises at that.

Dr. GREENHOUSE. I would like to be a little bit more precise in interpreting the question and in what I said. When I said, "Don't take something too seriously," I did not mean, "Don't take the clustering technique seriously," but rather that the clustering should be considered in a descriptive sense, not in a probabilistic sense. You form the groups of your individuals into neurotics, anxiety groups, reactive depressants, and so on, but you don't make the inference this technique would similarly classify other individuals in another sample in the same sense. That is what I meant by "not taking clustering techniques seriously." It means you go on to validate by other experiments if that is your problem, or else you go on to your next step in your experimentation or research.

Dr. FLEISS. I should also like to ask Dr. Gleser: In view of your down to earth approach to psychometrics, would you agree that the availability of computers is something of a mixed blessing? It seems what is done routinely is that data are gathered in the field, they are processed, key punched, fed into a computer, the results come out, and it is only at this last stage that investigators get to look at the data.

With respect to the application of computers to clustering, I think this has the danger Dr. Torgerson mentioned: It is possible in a clustering program indiscriminately to throw variables together, some of which do discriminate and some of which do not. Wouldn't a preliminary view of the data, say looking at the marginal distributions of each of the variables, permit you to throw out from the final clustering procedure those variables which don't discriminate? (You are shaking your head, "No.")

Dr. Lyerly, you criticize a method of Dr. Zubin and myself which assumes normal distribution within subgroups. I have a feeling psychometricians are not willing to assume normality in subgroups because

they see the raw data. The point is that it is possible to have distinct normal populations, but thrown together in such a way that the resulting distribution which we see before us is far from normal. The fact that a distribution is skewed does not remove the possibility that normal distributions lie underneath it; I think this should be borne in mind.

With respect to Dr. Rao's final point, I imagine your model there might very well be applicable to assigning newly admitted cases to one, two, or more diagnostic categories, especially when patients come in at different stages of their illnesses.

Dr. GLESER. I certainly agree with Dr. Fleiss that computers are likely to be a mixed blessing. I am afraid too many of our clinicians think all their problems will be solved by just putting enough numbers into a computer, and hence will have too much confidence in the results that come out. That is why I have tried to emphasize the relationship between what comes out and what sorts of assumptions are being made. Unless we have the assumptions the clinicians want, we are not going to get the right answers from computers.

Dr. LYERLY. I agree with Dr. Gleser that the phenomenon of marginal normal distributions does not preclude the identification of clusters. On the question of normality, many of us have been disturbed from time to time by some of its implications. Much of the psychometric data many of us deal with have been normalized through prior standardization on, hopefully, a relevant sample. These are in effect transformations of the original scale. That can be done in many cases, although there, again, we never know in advance what happens to the bivariate and multivariate distributions. It is a tricky problem, but I for one would prefer not to insist upon normality in anything. For example, among the variables that have been found to be most closely associated with susceptibility to and recovery from

mental illness in one way or another are such things as age, sex, education, and marital status, none of which is normally distributed. Yet we can and should make use of them whenever they will help us accomplish our purposes, whether they are normal or not.

Dr. RAO. I just want to make a very brief remark. The solutions which I give are applicable to cases where the mean characteristics of different states are linearly related. I also give a method of approach when the mean characteristics do not lie on the straight line, but depend in a known way. But the solutions are extremely nice and simple when the mean values of different states lie on straight lines.

Dr. HAMILTON. Dr. Torgerson expresses great doubt as to whether the notion that one illness per person is a useful assumption. This is, in fact, not only a useful assumption in practice, but it is the most useful assumption. When a patient gets an infectious fever, we generally assume the rash is due to, say, scarlet fever and not also to smallpox, measles, and one or two other things. I think it is very rare for anybody to be caught on this assumption of "one disorder."

When we deal with chronic disorders, we may easily be faced with, say, the problem that a given elderly patient may suffer not only from senile dementia, but from high blood pressure, rheumatoid arthritis, chronic duodenal ulcer, and a stricture of his urethra. In practice, however, he doesn't come for treatment because he is worried about all of these things, but because of his senile dementia or his stricture. He puts up with the others. So it is true that a patient can suffer from more than one mental disorder, but when one examines the history one finds one can distinguish it into that which has been going on for a long time and the particular problem which is worrying him at the moment and has brought him hot

foot to the psychiatrist. It is this with which one is concerned. So the assumption of one illness per person usually presents no difficulty in practice, because we are assigning these persons to a diagnosis. This, of course, is quite different, as was pointed out yesterday, from the problem of classification of disorders.

Nevertheless, it is true that the theoretician sometimes harps on a problem which the practical man has ignored or forgotten. As Dr. Rao has pointed out, the discriminate function should be sensitive to the difference between the two groups, but not sensitive to the stages or phases of a diagnostic group. This is absolutely fundamental, and it is very easily forgotten; I know, because I forgot it myself in the course of some of my work.

Dr. TORGERSON. To get back to the first question, "Is it a good idea for a person to have only one illness?", I think there is a difference here that shouldn't be overlooked or confused. It might make very good sense to diagnose a person having only one disorder, but when you are setting up an initial classification scheme you don't know what the variables are, and how they are sensitive to different disorders. You cannot ignore this fact. You can't, like the old man in Dr. Hamilton's example, decide this is the only thing that is important. I think it is a different problem.

Dr. HAMILTON. Dr. Torgerson, you show that different cross products yield different configurations. This is perfectly true, but I would like to point out that from the topological point of view these configurations are identical. We are not interested in the relative distances between configurations. What we want to know is the number.

Dr. TORGERSON. If we have more variables, we will come up with different clusters and different numbers of clusters, based on the index used. The clustering techniques are

not insensitive to these variables or these indices.

Dr. LUBIN. Most of the approaches in this discussion on methodology have been concerned with interdigitating the whole world. Some of the procedures described have taken into account very small groups, but for the most part we have been concerned with slicing the cake very neatly into large, even slices and reducing the number of crumbs. We ought to be paying attention to the crumbs; they have more consistency and homogeneity in the natural order of things.

Consider the problem of phenylketonuria. People appreciate now that phenylketonuria is a disorder which can be corrected through a carefully controlled diet. The phenylketonuric child was pulled out of a vast universe of mentally deficient children and has had his problem solved.

Another disorder which affects only a small number of children is the Kanner syndrome of infantile autism. I think there is sufficient statistical evidence to indicate infantile autism is a discrete disease; from the standpoint of decision theory the error in lumping a truly discrete disease with childhood schizophrenia is far greater than the opposite error of mistaking a subvariety for a separate disease entity.

Dr. GORDON. The question Dr. Rao addresses himself to in the use of discriminant function, that of assignment, would correspond to the art one learned as an intern of differential diagnosis. In psychiatry the process does not correspond with the discriminant function model which assigns some degree of probabilities of a function, but seems to be one of sequential categorization of a series of sequential decisions of a "yes" and "no" type. My question is whether or not there are alternative mathematical approaches to this process, and does he know of any attempts to test these alternative approaches in fields of medicine?

Dr. RAO. I have not fully understood the question. What I feel is that if there are well-established groups and if we have a certain set of measurements on a particular individual, the decision is of the type whether he belongs to one group or he belongs to the other group. If you are thinking of a sequential procedure, we allow for the possibility of not making a decision on the basis of a given set of measurements. If we are not able to decide, then we take certain additional measurements which together with the earlier ones may enable us to come to a decision. There is no statement of probability attached to a decision rule. The attempt is to minimize the number of wrong classifications in the long run or the loss which is expressed in any other way.

Dr. McQUITTY. I suggest we attempt to develop theories designed specifically to help improve measurement in the area of psychopathology.

I have attempted in my own work to apply the hypothesis that a problem is more apt to be solved if we develop theory in relation to the specific requirements of our problem. I have used a theory of types and developed this statistical definition: "Types are categories of people of such a nature that everyone in the category is more like every other person in the category than he is like any person in any other category."

This definition generates statistical methods. You then apply the statistical methods to data, and look at the results. This helps to revise the theory and to improve still further the assessment method.

A next logical step as a result of this conference is to attempt to develop theory which will generate data uniquely appropriate to the assessment of individual differences in psychopathology.

Dr. GREENHOUSE. It seems to me the danger in what you say is we would await theories on schizophrenia before measuring the variables these theories imply in

order to get classification schemes. There are four or five different schools of psychiatry, and I can easily see we would have four or five different theories. I suspect, worse than that, we might have a thousand different theories depending upon the individual researcher. I am not sure at this particular time this would lead to anything more scientific than has been going on thus far.

Dr. KRAMER. There was a statement in Dr. Gleser's paper I think may be at the nub of at least one of the problems of classification: "The need for new methods of classification of psychiatric illness is often supported by the fact the present diagnostic categories do not relate well to outcome with various therapies."

What is meant by "well" in this particular instance? For example, what result would one need in the outcome of therapy to conclude that the categories in the classification are satisfactory? I can demonstrate from diagnostic data we now have which can discriminate between classes of schizophrenics. We can subdivide schizophrenics into those who are married and those never married and demonstrate large differences in outcome. Within 6 months, 80 percent of married schizophrenics will be out of the hospital and only 50 percent of the "never marrieds" will be out of the hospital. Classification of schizophrenics by marital status discriminates between them in relation to outcome. Is there a presumption in Dr. Gleser's statement that if there were a better way of classifying schizophrenics, both married and single schizophrenics would have a better outcome? For example, a 100 percent improvement rate for both groups?

I would like to point out that even in homogeneous groups of patients used to evaluate therapy, there may or may not be a significant difference between experimentals and controls. In this instance, I am really at a loss to know how one can use the variable of outcome to decide whether

you have good or bad diagnostic classifications.

The basic problem revolves around the reliability and validity of the diagnosis. How can the observations be made so as to minimize observer error in arriving at a diagnosis? What invariant set of characteristics must diagnosticians observe in patients before assigning them to a given category? Are diagnosticians using a given term indiscriminately? The problem is not how to classify patients so that the outcome data will relate well to diagnosis, but what observations do we need in order to classify patients into existing diagnostic categories so that whatever the results with respect to outcome, the results reflect the true state of affairs.

Dr. GLESER. In the sense to which you refer, I believe I would say I used "well" here in terms of a useful differentiation.

In regard to your remark about married or single patients and their outcome, this is precisely the sort of thing I am talking about, except it happens to be only one dimension that is being used. But to the extent you find that noting whether the person is married or single helps you in predicting outcome, you have a classification that is useful. Now, if you were to find the married person with two children had a better prediction, you might also get a classification based on both of these characteristics, you see.

Dr. BALL. I would like to suggest maybe there is some way we can have our cake and crumbs, too. The idea of measurements is crucial. Obviously, you have to get good measurements. The next question is: How do you find them? Cluster-seeking techniques used cleverly seem to offer some hope, but cluster seeking by itself is not enough. You have to look at the details in the data. This means you have to have a computer, because nobody manipulates 20 dimensional variables easily without computational aids. You can make many kinds of measurements on the clusters. However,

they may not tell you very much about how the points within the cluster are distributed. Using the computer, you can generate a "picture" of a cluster.

Consider a situation in which a statistician and a clinician look at this picture of the data together. (The picture is on a TV tube.) They have the option of using the computer to provide factor analysis or principal components analysis or histograms of linear combinations of variables, or cluster-seeking techniques. The results of such analyses would be made available almost immediately and peripheral programs could be provided to aid in the manipulation of the output. The clinician may then say, "These results suggest an hypothesis I would like to propose for test." The statistician says, "Push three buttons and test it." And the statistician can be there to insure proper use of statistics. There is an interaction between the two of them. This type of presentation of data can be very powerful in looking at the details. And, by sharing this common experience of examining the data together, the statistician and the clinician should be able to communicate with each other more effectively.

Dr. FLEISS. How do you get a statistician and a clinician to sit down together?

Dr. BALL. You chain them. (Working together may also help make them see themselves as working on the same problem, although from different points of view.) I am not saying there is any magic in a computer. There is nothing but a lot of hard work. The computer may make some of the hard work a little easier for you.

Dr. BAHN. Dr. Rao's presentation and some others indicate the need for longitudinal observations on people. I think that has been somewhat overlooked in this discussion. It would be very helpful in typology.

Dr. GLUECK. Dr. Rao pointed out the different ways the term "classification" is

used. We have not fully defined the terms as to the semantics of classification. This has been a source of difficulty throughout our discussions here. It would serve a useful purpose if we could fully define terms like classification, grouping, and taxonomy. These are used in different ways, making discussion difficult.

How much attention are we paying to the dynamics of the conditions we discuss and how much to their statistics? Speaking in terms of mathematics, do we need a system of mathematics that will take care of movement, of changes in patients' conditions from one time to another—a special mathematics for that purpose—or is the mathematics we have perfectly adequate?

Dr. GREENHOUSE. Yes, we do need a mathematics to take care of changes over time. This is complicated because the changes may not be in one characteristic but in a multiplicity of characteristics. We require methods in multivariate time series.

Dr. KLEIN. There has been an argument between people who believe in a hypothetico-deductive approach and people who feel there should be some way of coming to structural theoretical conclusions inductively from the data. It is suggested clinicians and statisticians should work together closely, clinicians coming up with the hypotheses and statisticians testing them out. I agree. There were also wry queries from the audience, "How do you get them to do that?" This is an organizational problem, not one of individual personalities. In our present research setup we have not endowed organizations that systematically allow for this type of interaction.

Dr. CLYDE. I have a question for Professor Rao. On page 14,¹ at the end of the page, you say: "The value of c is determined

¹ Advance text sent conference members before meeting.

such that the level of significance has a given value." In determining the value c , do you assume independent chi squares?

Dr. RAO. I think the chi squares are dependent, and this should be taken into account in determining c .

THE CRITICAL REVIEW AND DISCUSSION OF THE METHODOLOGY INVOLVED IN A SERIES OF ATTEMPTS TO DEVELOP NEW TYPOLOGIES

A series of research reports are presented in which investigators describe their methods of development and their attempts to validate new systems for classifying the mental disorders. These investigations differ in the type of phenomena selected for study but all involve the analysis of multivariate data. The program includes reports both on the more focused descriptive and phenomenologic approaches to classification and on investigations which attempt to encompass variables from several sources, e.g., the behaviorial and the physiological, in the development of a system.

Each investigation is critically reviewed by a panel of discussants who then use the reports as background for appraising the current status of technical development and for exploring and clarifying basic substantive issues in the field.

The Phenomena of Depressions

*Roy R. Grinker, Sr., M.D.¹ and Jum
C. Nunnally, Ph. D.²*

Introduction

Experiences during and after World War II intensified an interest in psychodynamics by American psychiatrists leading to the development of stereotypes and a lessened concern with observations and description of clinical entities. On the other hand, the inadequacy of current nosological classifications stimulated the military psychiatrists to modify our systems of diagnosis (1). These efforts are continuing because of psychiatrists' resurgent interest in behavior. Studies on the therapeutic effectiveness of a hospital milieu and the efficacy of a wide variety of new drugs require rating scales for behavioral changes which cut across diagnostic categories. Clusters of traits became the source of classifications of types within the formal diagnostic categories.

When we became interested in the depressive syndrome, it seemed to have become the forgotten disease although about half the population in private psychiatric hospitals were admitted and discharged with this diagnosis. It is a self-limited cyclical affliction although each attack may last a long time without or even with treatment. The frequency, painfulness, and family disruption caused by depression stimulated little work with the syndrome and little has been added to its understanding since Kraepelin's and Bleuler's descrip-

tions which were only somewhat better than the accounts left by the ancients.

Applications of psychoanalytic theory and methods had defined stereotypes of intrapsychic dynamics during depressive attacks and within the premorbid personality. Precipitating factors were attributed to events corresponding to dynamic theory, and contents of communications were stressed. These advances in understanding did not help in adequately determining the indications and contraindications for the modern therapeutic approaches of psychotherapy, electric shock or antidepressive drugs. An unfortunate byproduct of focusing on the dynamics of depression has been the underemphasis on sound clinical observations and adequate descriptions of these and other mental patients. Most of American psychiatry is dynamic psychiatry and the word descriptive has become an appellation of derogation. As a result, the details of clinical syndromes are little known and the natural history of psychiatric diseases has been neglected.

When in 1954 we decided to take a "new look" at the depressive syndrome, etiological considerations had for long crystallized the classification of endogenous and reactive depression (2). Behavioral criteria maintained the diagnoses of neurotic and psychotic depressions. From these divisions nor from psychodynamic features could correlations be made with demographic data, physiological or biochemical variables, longitudinal course, recurrences, or therapeutic outcome. A new system of diag-

¹ Department of Psychiatry, Michael Reese Hospital (P.&P.I.), Chicago, Ill.

² Professor of Psychology, Vanderbilt University, Nashville, Tenn.

nostic classification of types of depression seemed to be badly needed.

Such a diagnostic typology is the basis of science which has been so successful in furthering advances in medicine. Empirically derived clinical types should furnish the basis for valid correlations with other systems as well as facilitate etiological and prognostic probability statements.

Purpose of the Research

The purpose of the research was to gain an understanding of some of the prominent trait-dimensions along which depressive patients vary with respect to their illness. Secondary purposes were to correlate those traits with background characteristics of patients and to explore some of the judgmental processes of psychiatrists with respect to depressive patients. Thus at the outset we had a "how" and a "why" as well as methodological questions.

Subjects

All subjects in the research were persons who had been diagnosed as depressed on admittance to the Psychiatric Institute of Michael Reese Hospital. With whatever fallibilities may have been entailed, the diagnosis was made on the basis of the conventional psychiatric interview. At this point, neither patients nor admitting physician had contact with our research activities, and none of our trait measures was employed in the diagnosis. The aim of our research was to investigate the characteristics of depressive parties after they had been so diagnosed by conventional psychiatric approaches, and for this purpose we were willing to take the full variety of whatever types of people are usually diagnosed as depressive. By this method we would probably include non-depressives who would serve as controls and exclude a few only. In all, 117 such patients participated in our research, 21 in a pilot study and 96 in the major studies.

Raters

Our studies were intimately involved with the measurement of behavioral characteristics from which later we might be able to develop more objective measures. It was necessary then to rely instead on human judgments regarding the behavioral characteristics of depressive patients. Those judgments were objectively recorded through the use of various rating methods, and they were objectively analyzed by statistical and mathematical methods; but nonetheless, one (feelings and concerns category) stemmed from the silent intuitive processes of the raters. The other (current behaviors) was derived from observations.

Since the research results hinge entirely on the particular raters employed in the studies, one naturally wonders to what extent the results can be generalized to other psychiatrists in other medical settings. A score of psychiatrists served as raters in different phases of the research. They ranged in age and experience from new residents to psychiatrists with many years of experience. It would be hard to estimate the effects on the research of whatever particular esprit there was in the institutional setting or of the background and personal characteristics of the psychiatrists. Some investigations were made of differences in our psychiatrists in reliability of ratings and in stereotypical conceptions of depressive patients. To our satisfaction two subsequent investigations made elsewhere have essentially confirmed our results.

Procedure

From start to finish, the intentions of the research were to (a) obtain a large sample of persons diagnosed as depressive, (b) have ratings made of those patients on a large number of particular traits thought to be related to depression, (c) statistically analyze the ratings in such a way as to arrive at more general traits or factors, and (d) cor-

relate those factors with other characteristics of the patients.

Traits

As much time probably was spent in gathering traits for use in the studies as was spent in the subsequent investigation of those traits. For this purpose, the major works on depression were studied, and from those were gleaned a very large number of particular attributes that had been ascribed to depressive patients, e.g., "feels hopeless," "dryness of the mouth," "dresses carelessly," "has dreams of suicide." In addition, many particular traits were suggested by the experiences of psychiatrists participating in the research. The criterion for including a particular trait was that it had been mentioned, observed, or reported by someone as possibly differentiating depressive patients from other types of psychiatric patients, from normal persons, or depressive patients from one another. At this stage the investigators were quite liberal in their inclusion of particular traits, because the main purposes of subsequent studies were to determine if those traits could discriminate depressive patients from one another.

A study of the traits suggested that they could be meaningfully divided into groups as follows:

1. *Precipitating experiences*: Events of recent history that may have pushed the patient over the brink into a depressive state, e.g., loss of loved one, major illness, economic loss.

2. *Dreams*, e.g., has frightening dreams, dreams of suicide (manifest content).

3. *Physical symptoms*, e.g., loss of weight, high blood pressure, breathes rapidly.

4. *Feelings and concerns*: The inner dispositions of the patient, which were necessarily inferred by the rater, e.g., feels a failure, concerned with death or dying, feels envious of others.

5. *Current behavior*: Either visible actions of patient (e.g., asks many questions, speech is accelerated) or traits that

require only a low level of inference on the part of the rater (e.g., tries to be witty and charming, denies need for help).

6. *Meanings of the precipitating events*: Interpreted by the investigators.

Because the research mainly was concerned with psychological and behavioral traits relating to depression, the major emphasis in our studies was on traits concerning "feelings and concerns" and "current behavior" (4 and 5 above). This emphasis was further strengthened by pilot studies of the other three classes of traits.

Rating tasks

Each trait in the first three classes above was rated as either present or absent in each patient. In pilot work with traits relating to "current behavior" and "feelings and concerns," the Q-sort was employed. Although results from the pilot study were quite suggestive, because of psychometric problems with employing Q-sorts on materials of these kinds, conventional rating scales were employed in the major studies. Each trait was rated on a seven-step scale, anchored by "not present" and "present to a marked extent."

In the pilot studies, raters were, in comparison to usual standards, well trained to perform the ratings. They worked closely together on formulating the lists of traits, on applying the trait lists to a small number of patients, and on deciding the future course of the research. Also, raters in the pilot study were, on the average, more experienced psychiatrists than those who participated in the later study. In the pilot study ratings were based mainly on unstructured and structured interviews with each patient observed through a one-way screen by another member of the team who made suggestions for questions to be asked in subsequent interviews.

In the major study, ratings were made by psychiatric residents on the ward of each patient. The residents had far less training in the use of the rating scales than was the

case for psychiatrists participating in the pilot study, but the residents had much more direct contact with patients. For reasons which will be mentioned later, there was much circumstantial evidence that ratings by residents based on ward contact were superior to those of the better trained raters in the pilot study.

Analysis and Results

Results from the pilot study made it apparent that there was little to be gained from continued study of dreams and precipitating experiences, at least not by the methods we were studying them. Psychiatrists reported almost no dreams, even though they realized that they would be required to make ratings of dreams. Results for precipitating experiences were highly scattered: few were noted for each patient, those did not appear to go together in any meaningful manner, and none was related to "current behavior" or "feelings and concerns" traits. From these findings, and our continued thinking, we concluded that one precipitating experience is not systematically related to the behavioral manifestations of depression. Whereas there may be definite precipitating experiences (plural and cumulative) in many cases, almost anything unfortunate will serve for that purpose, and, whatever it happens to be, it is not predictive of the behavioral characteristics of the depression itself.

Numerous physical symptoms were noted, and, in most cases those were ones that typically are identified with depressive states. A reduced list of physical symptoms then was carried over into the major study. Those were combined with current behavior traits to form 1 overall checklist of 139 items, each trait being rated as either present or absent. Since in the pilot work most of the current behavior traits occurred with moderate frequency, most of those traits were retained for the major studies.

Most of the remaining pilot work and the major studies concerned factor analysis in

one form or another. Except for one analysis to be discussed later, the analyses were of product-moment correlations. In some cases multiple-groups methods were employed to test hypotheses about supposed groupings of traits. In other instances the centroid method and the principal axes methods were employed as agents of discovery, the results of these being subject to rotation either by hand or by computerized analytic methods.

The first factor analysis was that of composite Q-sorts with the "feelings and concerns" traits, the composite sort for each of 21 patients being obtained by averaging the "sorts" by 4 psychiatrists. The three factors obtained from this analysis were helpful in composing a shortened list of traits for the feelings and concerns (from 111 to 47) rating form used in the major studies. Results discussed in the following sections will be from the major studies, in which the feelings and concerns rating form and the current behavior checklist (including physical symptoms) were applied to 96 depressive patients.

Reliability

Because ratings notoriously have only modest reliability, it was necessary for us to make extensive studies of that issue. In the pilot study it was found that psychiatrists did not agree very well with one another. Surprisingly there was more agreement on "feelings and concerns" than on "current behavior." In other words, psychiatrists agreed more when it was necessary to make an inference about inner states of the patient than when the behavior was available to the naked eye. This told us, we think, something about the habits of psychiatrists in interview situations.³

³ This finding has had a profound effect on subsequent research still in progress by Grinker, Werble and Drye on the borderline states in that observations and descriptions of behavior are being made by the entire body of attendants in a nursing unit around the clock, dictated in detail, and the resultant protocol rated by trained, skilled professionals.

Because of differences in rating tasks in the pilot study and in the major study, it was hard to make a direct comparison of the reliability of ratings in the two, but there was much circumstantial evidence that reliability was considerably higher in the latter. This was so in spite of the fact that ratings in the pilot study were by more experienced psychiatrists, more experienced professionally and more experienced in the purposes of the study. This probably echoes what has been found before about the reliability (and perhaps the validity) of ratings: the professional standing of a rater is not nearly as important as the amount and kinds of information available about the person to be rated. The psychiatric residents who made ratings in the major studies had considerable daily contact with patients in lifelike situations.

Factors among feelings and concerns

Five (rather strong) factors were found in the "feelings and concerns" ratings. There were some surprises, in that in a number of cases items loaded highly on factors where one ordinarily would not have anticipated their presence. Loadings on the factors are higher than those typically found for rating scales. Internal consistency reliabilities for groups of items used to define the factors ranged from 0.61 to 0.90.

We have established the existence of five factors, which are patterns of traits descriptive of the feelings and concerns of these patients. These factors illustrate aspects of patients, and, although they indicate what may be predominate for some, any single patient may show evidence of more than one factor. They are not mutually exclusive; nor do they attempt to illustrate all aspects of any single patient. These factors may be roughly characterized as follows:

(I) A factor describing characteristics of hopelessness, helplessness, failure, sadness, unworthiness, guilt, and internal suffering. There is self-concept of "badness."

(II) A factor describing characteristics of concern with material loss and an inner con-

viction that this feeling state (and the illness) could be changed if only the outside world would provide something.

(III) A factor describing characteristics of guilt over wrongdoing, wishes to make restitution, and a feeling that the illness was brought on by the patient himself and is deserved.

(IV) A factor describing characteristics of free anxiety.

(V) A factor describing characteristics of envy, loneliness, martyred affliction, secondary gain with gratification from the illness, and attempts by provoking guilt to force the world into making redress.

The clinical interpretation of these factors suggests that factor I is the essence of depression, and hence its strength indicates the depth of the affective disturbance. The anxiety factor, IV, seems to indicate activity in the process and perhaps also a signal of mobilizing or declining unconscious aggression. On the other hand, the remaining factors indicate varying attempts at defense and resolution of the depression. Hence factor II indicates the projective defense; III, the restitution resolution; and V, the attempt by enslavement of external objects to deny anger, and secondarily to regain love. The control groups seem to show that the diagnosis of the depressive syndrome is contingent not only on the depressed affect but also on the presence of anxiety. In fact, in the presence of minimum sadness, anxiety is enough to weigh heavily for the diagnosis of depression. Finally, the nondepressed patient whose admission diagnosis was accurate has in common with depressions the factor of dependency and demand for secondary gain.

Factors in current behavior

Ten factors were found in the "current behavior checklist." Because of their number, and because their meanings are not all obvious, no efforts will be made to give capsule interpretations of them. After the fact they make sense (as most factors do), but we could not guess

their nature in advance, and on seeing them no one is likely to say, "Just as I suspected." Also, on inspecting those factors one gets the distinct impression that they would be useful in describing any psychiatric patients, depressive patients included, indicating that there are a limited number of behavioral "final common pathways," whereas the factors found in the feelings and concerns traits appear more specifically relevant to depressed patients. Internal consistency reliabilities for the groups of items used to define the factors ranged from 0.64 to 0.86.

These factors tend to be less sharp and distinct than the feelings and concerns factors, which reflect our finding that behavior is an area of less interest to psychiatrists and that our behavioral observations are not as accurate as our observations of content. Like the other factors, they are not mutually exclusive, and any single patient participates in a number of factors. These behavioral factors may be roughly characterized as follows:

(1) Characteristics of isolation, withdrawal, and apathy; (2) characteristics of retardation, slowing of thought processes and speech, with little regard for personal appearance; (3) characteristics of general retardation in behavior and gait, but less isolated and withdrawn than factor (1); (4) characteristics of angry, provocative, complaining behavior; (5) characterized by somatic complaints, including dizzy spells and constipation; (6) characteristics which sound like an "organic" syndrome: impairment of memory, confusion, inability to concentrate, and limited and repetitive thought content; (7) characteristics of agitation, tremulousness, and restlessness; (8) characteristics of rigidity and psychomotor retardation; (9) characterized by somatic symptoms such as dry skin and hair, along with some abnormalities on physical examination; (10) characteristics of ingratiating behavior, attempts to help patients and staff, appreciative for the interest of the staff and the facilities of the hospital.

Relations between two sets of factors

After factor scores were obtained for the 96 patients on the 15 factors, correlations were computed among the factors. We were surprised to find that feelings and concerns factors did not correlate at all with current behavior factors. For example, anxiety did not correlate with agitation. This suggests that, to the extent that one can consider such ratings of feelings and concerns as valid indicators of the "internal dynamics" of depression, such dynamics are not related in any known systematic way to the patient's outward behavior. It appeared as though the outward signs were not really symptomatic of the depression itself but, rather, that they were expressions in a characteristic personal behavioral pattern of an inner distress.

Correlations of factors with background variables

The factors should correlate with something in the history of the patient, and, more important, they should correlate with what happens to the patient in the course of treatment and beyond. We were able only to make a few exploratory investigations of this kind. We found interesting differences on our factors between males and females, e.g., the former tend to focus on external concerns and the latter are dominated by guilt. We found that older patients tended to be higher on all of the current behavior factors, showing that older people are more prone to show the aggravated outward signs of depression.

One of the most interesting findings with regard to the patient's stay in the hospital concerned changes in diagnosis. We found a regular relationship between changes in diagnosis and our factor of anxiety. Patients who originally were classified as depressive but later were classified otherwise were rated originally as being high on anxiety but not high on our factor of "dismal affect." The reverse was true of patients who were not diagnosed as depressed originally but were

diagnosed as depressed after prolonged hospitalization. It has been confirmed elsewhere that anxiety is not only an indicator of activity but also of elevated adrenocorticoid steroids, all of which are related to intense suicidal drives.

Stereotypes

At times in our studies we felt that we were learning as much about psychiatrists as about patients. Studies mentioned previously about the amounts of agreement about psychiatrists working with different types of information about patients was instructive in that regard. Special studies of psychiatric stereotypes provided other suggestive evidence. Using our rating scales, psychiatrists were asked to describe the typical depressed patient. Those ratings were factor analyzed (correlating psychiatrists with one another), and the loadings of psychiatrists on the factors were correlated with various aspects of their previous ratings of patients. We found, for example, that more experienced psychiatrists had stronger stereotypes than less experienced psychiatrists. Also, we found that the relative similarity of stereotypes for two psychiatrists was somewhat predictive of the amount of their agreement in rating particular patients. In studying the major factor in the stereotype it was obvious that "it" was a woman, a separate factor (shared by only some of the psychiatrists) relating to characteristics of male depressives.

Profile clusters

One of the final analyses in our studies was to investigate clusters of patients in terms of our 15 factors. Each patient was scored on each of the factors, sums of cross-products of those factor scores were obtained for all possible pairs of patients, and the resulting 96 x 96 matrix of cross-products terms were submitted to a centroid analysis. Essentially what was different about this analysis from more customary approaches to factoring is that it was a "raw score" analysis rather than the more custo-

mary analysis of correlations or covariances (3). The analysis produced factors, each factor relating to a type of depressive patient. The major factor was interpreted as the "garden variety" depressive, and the other two factors were interpreted as less prevalent types. Factor patterns were developed from the combination of 15 factors of both trait lists. As a result four factor patterns were elicited from which clinical profiles can be described to serve as fairly sharp hypotheses for future testing. The factor patterns are as follows:

(A) Feelings: dismal, hopeless, loss of self-esteem, slight guilt feelings. Behavior: isolated, withdrawn, apathetic, speech and thinking slowed with some cognitive disturbances.

(B) Feelings: hopeless with low self-esteem, considerable guilt feelings, much anxiety. Behavior: agitation and clinging demands for attention.

(C) Feelings: abandonment and loss of love. Behavior: agitated, demanding, hypochondriacal.

(D) Feelings: gloom, hopelessness, and anxiety. Behavior: demanding, angry, provocative.

Implications for Diagnosis

The major purpose of our studies was to provide some understanding of the traits of depressive patients. To achieve that understanding, however, we were mainly concerned with central tendencies—central tendencies of all our depressive patients and differences in central tendencies of subgroups of our patients. Of course, even if one finds marked differences in such central tendencies, that does not mean that differences could be used effectively for diagnosis. In diagnosis one must take account not only of differences in central tendency but also of variabilities within groups. For example, as a group our depressive patients were moderately high on a factor relating to anxiety, and this helps us understand the modal characteristics of depressive patients,

But our depressive patients varied markedly with respect to anxiety, so much so that it would provide only a small amount of discriminative information with respect to normal people. Similarly for the various subgroups determined by our factors: Such subgroups differ markedly in central tendencies with respect to the factors, but the variabilities within groups also are large.

Use of our factors in diagnosis

It is possible that our 15 factors will prove useful in diagnosis, but, if so, there is psychometric work yet to be done with them. The reliabilities of some of the factors are too low to permit efficient diagnosis with them. More items should be added to some of the factors, and new studies should be done to determine if the new items are placed where they should be. In addition, the factor measures could probably be improved by clarifying the instructions and experimenting with different types of rating scales.

The meaning of diagnosis

The problem of diagnosis is not so much a mathematical problem as it is a psychological problem. It is proper to ask "Diagnosis for what?" If diagnosis is to be more than a legal obligation, or make-work activity, then it should be relevant to the prognosis of the illness or, even more to the point, predictive of the effects of different treatments. In other words, any set of diagnostic categories should have to prove itself in terms of its ability to add valid predictive variance to that which can be obtained from less formal ways of making decisions.

Presently, the major problem is that no one can say what class of decisions is most relevant to a scheme of diagnosis. Should these decisions relate to (a) reactions to different types of hospital environments, (b) effects of different types of drugs, or (c) improvement in different types of psychotherapy? The diagnostic categories that might be valid with respect to one of these

might be nearly worthless with respect to another, and the traits that might be required for forming diagnostic categories for one of these purposes might be different for the others.

One gets the impression that much of current diagnostic practice is not so much related to decisions about future events but is used to justify decisions that already have been made about patients. Thus, in justifying why a patient should be hospitalized, the psychiatrist can argue that this is because he is schizophrenic. In justifying why shock treatment was given to a particular patient, the psychiatrist can argue that this is because he is depressive. In justifying why a patient is allowed to go home, the psychiatrist can say that this is because the patient has a reactive depression and thus it can be expected to be transitory. In these instances, the impression is that the diagnosis was not so much instrumental in making the decisions but, rather, that the diagnosis was as much influenced by the decision as vice versa.

Another, and perhaps the most important, use of diagnosis is simply to provide a description of the patient with respect to important traits. If nothing else, this gives one some cognitive closure in dealing with patients, and, in an informal way, it probably plays an important part in daily interactions between psychiatrists and patients. In these interactions numerous decisions are made, e.g., to permit a patient to spend a weekend at home, but these decisions are so numerous and sundry, and different for different patients, that it is hard to see how they could be made the basis for an efficient scheme of diagnosis.

Typologies

In thinking of diagnostic schemes it is difficult to keep from thinking in terms of types. Methods for clustering profiles essentially search for types, and discriminatory analysis is concerned with differentiating among a priori types. Unfortunately, people are not sympathetic to the mathe-

mathematical requirements for employing such types, because most people prove to be atypical. This is evidenced in the major studies with the multiple discriminant function to date, where it has uniformly been found that there is far more overlap than separation among groups. Is there any reason to believe that mental patients will neatly divide themselves into "types" where the between-groups variance is large relative to the total variance?

It is however, difficult to keep from thinking in terms of types. We found it difficult to discuss one of our factors without thinking of how it would combine with other factors in an actual person. For example, we wanted to talk about the melancholic-anxious-guilt ridden patient. But to take account of all of our factors would require 14 hyphens, and long before that point is reached, meaning would be gone from the system.

Salient attributes

What usually occurs in actual diagnosis, and what may have much to recommend itself, is that a patient is described only in terms of salient attributes, and such attributes may be salient in two ways. First, attributes may be salient in that they are extreme in a particular patient. For example, if a patient is extreme only with respect to two of our factors, e.g., melancholic and anxious, then those two terms would offer a capsule description of the outstanding features of the patient.

Saliency in the above sense probably has been the basis for the presently employed diagnostic categories. For example, even though there may be other things wrong with him, if a patient is severely melancholic, he probably would be classified as depressed. If the most outstanding feature of another patient is disrupted thought processes, he probably would be classified as schizophrenic. Of course, atypicality alone is not the only aspect of saliency that is involved in diagnosis: The severity of the symptoms also are considered. Thus, if

a person were equally extreme in terms of melancholia and disrupted thought processes, he would be classified in terms of the most dangerous symptoms and declared schizophrenic.

Attributes may be salient in another way; some of them may be much more predictive than others of the outcomes of certain classes of decisions. Thus, for example, even if we obtained 15 factors, we suspect that one of them, anxiety, is more predictive than the other of the course of the illness and the response of the patient to treatment. Saliency in this sense relates to multiple-regression models for predicting criterion variables, and this approach to diagnosis has much to recommend it. It is not bogged down in the psychology of types and it can be employed quite flexibly with respect to any forms of treatments where the outcomes are measurable.

Mathematical models for diagnosis

In this section we will admit the future possibility of a computerized diagnosis-for-action and speculate about the most effective mathematical models for that purpose. The word "diagnosis" is so frequently linked with the word "classification," and the latter is so frequently mentioned with respect to discriminatory analysis, that it is easy to assume that the problem of diagnosis can be handled by the multiple discriminant function. If that type of analysis is employed in the usual way, it may not be the most effective approach. By "usual way" is meant that each person is placed in one of a limited number of categories and some utility is placed on the person's correct classification. This would be the case, for example, if each person is to be assigned to one of three drug groups. In that instance, it probably would be the case that many of the patients would fare much the same on any of the drugs and some of the patients would do very badly on any one of the three. Then if one must assign each patient to one of the three drug groups, it might be necessary to make some very bad decisions. To

take an extreme example, it may be that the probability is high that any of the three drugs will kill the patient, but, since the probability is only 0.8 that he will be killed by *A* rather than 0.9 for the other two drugs, he is administered drug *A*.

One way to get around the problems inherent in classifying patients in terms of discriminatory analysis is to include a no-treatment group. Then, unless the person has a high probability of fitting in one of the other groups, he would be assigned to the no-treatment group. But in doing this one is, in essence, shifting from the logic of discriminatory analysis to the logic of regression analysis. In discriminatory analysis, each person is assigned to the group where he will perform better (no matter how good that is in absolute sense); in regression analysis an individual need be assigned to a treatment only if the predicted outcome is good in an absolute sense.

Multiple regression analysis probably is a more effective paradigm for diagnosis than the usual forms of discriminatory analysis for another reason. Discriminatory analysis (and related methods of classification such as the use of centours) starts with a priori groupings of people, and the effectiveness of the analysis can be no better than the original methods for grouping people. Of course the present problem is that there are no effective ways of forming a priori groups of patients, and thus there is no way of effectively employing discriminatory analysis and the attendant methods of classification.

One possible way to obtain important groupings of patients is through methods for clustering profiles, such as that performed in one of our studies. However, by this method groups are defined in a circular sense in terms of behavioral characteristics, and, although this may aid in understanding the present characteristics of patients it is not obviously relevant for the outcomes of decisions about them.

One way to obtain groups for discriminatory analysis would be by experimenting

with treatments. Persons who respond much better to one treatment than another would be placed in groups with respect to those treatments. Discriminatory analysis could be used to differentiate those groups in terms of psychological traits, and, subsequently the obtained discriminants could be used for classification of new persons. In order to perform the necessary experiments, however, it would be required either to assign each patient to all treatments or to randomly assign each patient to one treatment, neither of which is practicable. In a sense, multiple regression would be hobbled with the same problem, but, with that method, corrections for restrictions in ranges would lessen the problem considerably.

If, and when, diagnosis-for-action is computerized, it probably will be based on the logic of multiple regression. In addition to its other virtues, it is free from questions of typologies and not dependent on the assertion of a priori groups. Although such an approach might be effective in a statistical sense, it would provide little help in understanding individual patients, for which purpose psychiatrists are likely to continue thinking in terms of salient attributes.

REFERENCES

- (1) GRINKER, R. R., Sr., and SPIEGEL, J. P.: "Men Under Stress," Philadelphia, Blakiston, 1945.
- (2) GRINKER, R. R., Sr., MILLER, J., SABSHIN, M., NUNN, R., and NUNNALLY, J. C.: "The Phenomena of Depressions." New York, Hoeber, 1961.
- (3) NUNNALLY, J. C.: "The Analysis of Profile Data." *Psychol. Bull.*, 59:311, 1962.

Supplementary Comments by Dr. Grinker as Presented at the Conference

I shall use this brief period not to repeat or to amplify our circulated paper, but (1) to present reasons for our clinical assumptions and approaches and (2) to indicate

how the weakest part of our previous research design has been modified and hopefully improved in subsequent investigations. It would be banal to say to this audience that what statistical analysis can put out is dependent upon what the clinician puts in from his own natural computer system, but it is not banal to say that his computer machines require continual calibration.

The practical questions concerning the results of treatment of various kinds seemed to have presented no great concern to clinicians until the advent of the new tranquilizing and antidepressive drugs. Each psychiatrist usually diagnosed, treated, and evaluated his results on few patients—hardly a scientific project. When the new drugs appeared a rash of rating scales were developed to cope with behavioral changes during clinical trials with drugs.

The so-called dynamic psychiatrist, however, has little interest in diagnosis or classification of various types of behavior. Concerned with the individual, he is more interested in his patients' problems and dynamics. He cannot be expected to contribute to a nosological classification. When he does, such abortions as "passive aggressive personality" have been induced.

An unfortunate concomitant has been the underemphasis on sound observation and descriptions which are considered obsolete and old fashioned. The APA report on "Training the Psychiatrist to Meet Changing Needs" published in 1963 considers observation, description, and classification as a biological technique in which the patient is treated as an object and as the first phase of the development of psychiatry as a science.

I find no need to quibble as to whether this is a medical or social science model. All science in general is concerned with behaviors. The argument yesterday, I felt, regressed to the either/or level, but each approach separately and/or together can be of value.

Behavior represents in actuality functions allocated to an hypothetical ego which filters perceptions on the one hand and actions on the other, which express reportable motivations, affects, defenses, and compromises, which employs symptoms and sublimations and demonstrates integrative capacities and disintegrative trends. Behavior is the final common pathway of all these processes.

In using the term "ego" I designate a process or an allocated function which has "structure" in the sense of an enduring pattern over time. Behavior as I use it has no connection with behaviorism as a theoretical position à la Watson and Skinner which depreciates the mental as an object of scientific investigation. Instead, behavior becomes a process for scientific investigation for which objective data are necessary to be acquired by special methods. From this hard data, speculations, even inferences, hypothesis building may follow ad lib.

Behavior can be observed, described, classified, quantified, modified, and analyzed within the matrix of scientific inquiry. All of the derived data are public and therefore can be coded, rated, repeated, replicated, and tracked through time.

The approach to behavior need not be contaminated by therapeutic interferences which modify behavior in an uncontrolled fashion and restrict, by virtue of the supposed ethics of therapists, the extent of observations. Behavior as the final common pathway of a wide variety of processes, actions, and reactions demonstrates not only liabilities or "illness", but also assets which are observed in the therapeutic dyad only with great difficulty.

As I have stated many times, we need to observe not only goal-seeking behavior as the repetitive neurotic trends or at the worst the symptoms of illness, but also the goal-changing activities of adaptation and creativity, indicative of autonomy of the ego and freedom to change the human situation. Such goal-changing activities

may be adaptive and creative as they place a burden on internal homeostasis. They may be developmental and related to crises in growth or they may be disruptive to the environment.

If the so-called third phase of social psychiatry suggested by the GAP report is to develop, its data must be from behavior as it is influenced by and influences the social community. Humans are not accepted or rejected by their society because they feel badly nor does the average person usually indicate his need for help because he is not happy enough. He becomes aberrant and feels differently because his behavior is not adaptive within a particular sociocultural environment.

Furthermore, we are beginning to realize that the therapeutic effectiveness in any form of psychotherapy is not from insight or at least from insight alone. Action and experimental behavior are the processes by which the validity of insight is learned. There is growing evidence that therapeutic effectiveness may be achieved by first modifying behavior which is subsequently followed by internal reorganization and sometimes insight.

All this indicates to me that psychiatrists should again learn how to observe and describe. The scientifically curious will classify behaviors into types and correlate them with a variety of biological, psychological, and environmental factors in the field within which behaviors transact with other behaviors.

Our investigations on depressions revealed that our two trait lists had different degrees of reliability. Inferences about "feelings and concerns" were highly reliable even though they did not conform to the usual stereotypes. When behavior was rated by observations over time by the unit resident and head nurse, reliability was high but not as good as for "feelings and concerns."

Our design for the observation, description, and rating of behavioral traits in depression included the typical method

utilized when specified traits are presupposed to be present or absent and their quantities can be estimated or measured. All other behaviors are then ignored as if they did not exist. As a result we are rating only a segment of behavior presumed to be salient, to the neglect of all others.

In changing this method I am reminded of what my mother used to say when I came to the dinner table after an afternoon of play, that my eyes were bigger than my stomach, and it is true that we have a tendency to include too many traits in our descriptions of behavior. Many traits, however, will be discarded later, for good reasons, based on theory which at the same time is being tested. We have to later regurgitate what we cannot digest.

The technique in the depression research also placed in the hands of so-called trained or experienced professionals or subprofessionals the collection of data as well as their ratings. Although this may result in high reliability it also contaminates the research with some bias.

In a new program of research on the "borderline" category, conducted with Drs. Werble and Drye at the Illinois State Psychiatric Institute, we are dealing with the borderline. Unstructured behavioral observations were made by the entire range of professional and subprofessional staff which serve the patient day and night in the course of their usual practice. The number was usually 15 to 18 observers who were not contaminated by any knowledge of our interest or meaning of the research. Thus a large number of observations over a long period of time (2 weeks) were made by staff intimately involved with the patients' daily life, and the many observers of the same incidents washed out individual bias.

Then each observer was interviewed by a skilled worker to elicit and tape record the behavior, including the verbal, data of all the patients' transactions within a known milieu. No inferences were considered data, only concrete behaviors and

statements to answer the question, "What did the patient do or say" in relation to all situations and all people. There was no feedback among the reporters, the minimum interventions on the patient were made, and the patient did not know he was being studied.

The setting is in a specific nursing unit with a stable staff. Observations are made 2 weeks after admission when the patient has settled down and 2 weeks before discharge, which is a universal stress stimulus. Followups from 1 to 5 years after discharge are being made and the ultimate condition assessed. The families are being rated as systems to determine what these patients experienced in early life and to what they returned.

The transactional verbatim protocols were then translated into clinical terms using 200 questions in an ego-functions framework. These were divided into five major categories with numerous subdivisions: (1) Outward behavior, (2) perception, (3) messages, (4) affects and defenses, and (5) synthesis, which enables us to make predictive statements to be tested by followup.

The complex rating procedure necessitated the production of a manual describing each category and defining each degree (from zero to nine) according to operational definitions and concrete illustrations. Two well-trained and reliable professionals completely outside the experimental setting rated each protocol and when more than two points of difference occurred two other raters reconciled the results.

Finally, the ratings of 15 to 18 protocols of each patient were combined into a single rating for each item. It is expected that the statistical studies now in progress will establish a clinical definition of the borderline syndrome and its subcategories.

Finally, I should like to emphasize that the objective data can be assessed in terms of ego functions as an index of mentation not typically observed in a dyadic relation-

ship and are an indirect measurement of the quality and quantity of internal psychological functions, but always with external reference. We utilized concept of ego in terms of adaptational theory when psychoanalysis returned to its position as an open system. This offers a hope that basic patterns rather than symptoms for disease concepts may be the basis on which systems of classification can be developed.

Our dynamic interpretation of the meaning of the five factor patterns in depression is a good example of how we can hypothesize (1) the essence of depression; (2) the indications and meaning of changes in the quantity of anxiety and (3) the defenses and restitutions.

In brief, I believe that behavioral transactional researches constitute a valuable bridge between internal psychological and biological processes on the one hand and social, cultural, and ecological factors on the other hand.

A Typology for Functional Psychotics¹

*Maurice Lorr, Ph. D.*²

Background

The aim of this paper is to present a digest of four studies,¹ each designed to identify mutually exclusive and homogeneous subgroups among psychotic patients. Emphasis will be on substantive findings rather than methods used.

Before launching into a description of these investigations and their findings, some of the methodological problems in-

¹ Much of what is reported here is based on a study entitled "An Approach to the Classification of Psychotics" (NIMH Research Grant 06653) conducted by M. Lorr, C. J. Klett, and D. M. McNair.

² Professor of Psychology, Department of Psychology and Psychiatry, Catholic University, Washington, D.C.

volved will be sketched. The view of the author is that the general problem is one of sorting people into subgroups or types when these are not known a priori. Similar problems are encountered by the zoologist, the political scientist, the dress manufacturer, and the occupational specialist. First it is necessary to choose and measure a set of characteristics on the basis of which individuals can be compared. Next some index of similarity between individuals must be selected. Then a procedure must be found for identifying the subgroups in the nonrandom population studied. Once the K subgroups have been distinguished, the next problem is to draw boundaries between them so as to separate them optimally. Finally decision functions must be specified for assigning new individuals to the subgroups identified.

There are certain conditions that must be met by the set of defining characteristics selected in a domain of interest. It is essential that all major sources of behavior variation be represented. Should some important sources of trait variation be excluded then the profiles delineated will be incomplete. The addition of new dimensions could alter these profiles in significant ways. A related requirement is that the defining variables be descriptive of independent dimensions. A major reason is that most indices of similarity assume that the profile elements are independent. One approach to the satisfaction of this condition is to conduct a series of factor analyses in order to isolate the major dimensions of a given universe.

Once the defining characteristics have been selected it is necessary to choose some index of resemblance between individuals. A wide variety of indices of similarity have been proposed. As early as 1898 Heincke suggested the use of a distance measure (D^2). Pearson (1926) proposed the coefficient of racial likeness and Mahalanobis described a generalized distance function. Cronbach and Gleser (1963) have reviewed most of these methods for assessing similarity. They

have shown that most coefficients may be subsumed under a general measure of distance between two individuals.

The third step in the development of a classification scheme is to identify subgroups or hierarchies of subgroups in nonrandom populations. The grouping process begins with a basic data matrix consisting of K scores on N persons (entities). The data matrix is then converted into a symmetric derived matrix of similarity indices among the N persons. The problem is to recover information about the grouping of persons given such a matrix.

Two formal algebraic models have been suggested for use in identifying homogeneous subgroups. Stephenson (1952) and Nunnally (1962) have recommended factor analysis as a solution. Others have suggested the use of Lazarfeld's latent class model. A variety of numerical techniques for defining groups called cluster analysis has an equally long history beginning with Zubin (1938). One of the earliest reviews of cluster search methods was by Cattell (1944). Since then McQuitty (1954, 1964) has proposed a considerable array of specific procedures such as linkage and typal analysis. Thorndike (1953) developed a clustering procedure later modified by Sawrey, Keller, and Conger (1960). Zubin, Fleiss, and Burdick (1963) have described a process resembling Thorndike's. Saunders and Schucman (1962) developed a procedure called syndrome analysis based on sets of profiles that are mutually closest. A detailed review of these and other procedures may be found elsewhere (Lorr, 1965).

The Measuring Instrument

The studies to be reported all utilized the Inpatient Multidimensional Psychiatric Scale (IMPS). The rating schedule is designed to measure 10 psychotic syndromes confirmed in repeated factor analyses (Lorr, Klett, and McNair, 1963). IMPS consists of 75 brief rating scales and dichotomous items. Ratings are made following interviews of 30 to 45 minutes in length.

Each syndrome is defined by a set of scales, from 5 to 11 in number, that measure a unitary pattern of behavior. The labels given the syndromes are intended to describe the underlying response tendency. For convenience in referral each syndrome is also identified by three letters. The 10 syndromes are briefly characterized below:

1. Excitement (EXC): The patient's speech is hurried, loud, and difficult to stop. His mood level and self-esteem are elevated and his emotional expression tends to be unrestrained or histrionic. He is also likely to exhibit controlling or dominant behavior.

2. Hostile belligerence (HOS): The patient's attitude toward others is one of disdain and moroseness. He is likely to manifest much hostility, resentment and a complaining bitterness. His difficulties and failures tend to be blamed on others.

3. Paranoid projection (PAR): The patient gives evidence of fixed beliefs that attribute a hostile, persecuting, and controlling intent to others around him.

4. Grandiose expansiveness (GRN): The patient's attitude toward others is one of superiority. He exhibits fixed beliefs that he possesses unusual powers. He reports divine missions and may identify himself with well-known or historical personalities.

5. Perceptual distortions (PCP): The patient reports hallucinations (voices and visions) that threaten, accuse, or demand.

6. Anxious intropunitiveness (INP): The patient reports vague apprehension as well as specific anxieties. His attitudes toward himself are disparaging. He is also prone to report feelings of guilt and remorse for real and imagined faults. The underlying mood is typically dysphoric.

7. Retardation and apathy (RTD): The patient's speech, ideation, and motor activity are delayed, slowed, or blocked. In addition he is likely to manifest apathy and disinterest in the future.

8. Disorientation (DIS): The patient's orientation with respect to time, place, and

season is defective. He may show failure to recognize others around him.

9. Motor disturbances (MTR): The patient assumes and maintains bizarre postures and he makes repetitive facial and body movements.

10. Conceptual disorganization (CNP): Disturbances in the patient's stream of thought are manifested in irrelevant, incoherent, and rambling speech. Repetition of stereotyped phrases and coining of new words are also common.

The syndromes have been shown to be sufficiently reliable in a variety of settings (Lorr, Klett, McNair, and Lasky, 1962). Some of the evidence for the validity of the syndromes has been reported (Lorr et al, 1963). In brief, it has been shown that the syndrome scores are related to ward assignment, to prominent features of psychopathology, and to conventional diagnostic class membership. The syndromes have also been found to be sensitive to change with tranquilizer treatment and to psychopathology as measured by other measuring devices.

Method of Analysis

The findings from four studies will be summarized here. The first investigation sought to find the psychotic subgroups among a random sample of acutely disturbed psychotics newly admitted to 16 State and university hospitals. The second and third analyses dealt with: (a) relatively chronic but newly admitted male psychotics, and (b) a sample of long-term patients examined in connection with two drug discontinuation studies. The fourth analysis was based on a nine-hospital NIMH study of a sample of newly admitted schizophrenics.

The methods of analysis were essentially the same in the first three studies described. Since the computer program developed for clustering profiles into homogeneous subgroups is capable of handling not more than 150 profiles simultaneously, it was necessary to subdivide each study sample. Thus each study sample was partitioned into three

subsamples. Next the clustering program was applied separately to the correlations (Q) and the congruency coefficients (C) among the 150 cases represented in each of the three subsamples. The clusters were first matched within each subsample across the two indices of similarity. Next the clusters were matched across the three subsamples. Finally a stratified sample was selected as a further check on the cross-sample and cross-index matching. Each clearly defined type found in at least two subsamples was represented in the stratified sample in proportion to its relative frequency. Analysis of the stratified sample provided not only a check on the accuracy of the matching but provided a test of cluster invariance. Any type or class evolved should be replicable under changes in the sample of persons examined providing it is represented in sufficient numbers.

Matching within a subsample across similarity indices involved several steps. Two clusters defined by a high proportion of identical cases were judged identical. In addition the 10 standard scores of cluster members were averaged. Then congruency coefficients between the mean syndrome profiles of the two sets of clusters being compared were computed. Clusters were considered identical if the congruency coefficient between their profiles was at least 0.75.

Type matching across subsamples was also achieved by comparing the congruency coefficients between the mean syndrome profiles of the various clusters. Two clusters were judged identical if the congruency coefficient between their profiles was 0.75 or greater and all other indices of similarity were negative or close to zero. Such matching was checked by resort to the clusters identified in the stratified sample. When subsample clusters had been correctly matched their representatives in the stratified sample were nearly all found in the same cluster evolved. Thus the matching by congruency coefficient could readily be confirmed or rejected.

Method of Labeling Types

The profile of a type can be characterized in terms of those mean syndrome scores that lie at some point above or below the general mean on the standard score scale. Since the syndromes are descriptive of behavior deviation, the higher the score the greater the deviation from the norm. Thus the most critical syndromes are those in which all members score above the mean. Syndromes on which all members of a type score below the mean are also important; they indicate the relative absence of pathology. On the other hand, syndromes on which members of a type score as frequently above as below the mean are simply undifferentiating and can be ignored.

For convenience and brevity let each syndrome be identified by a number as follows:

1. Excitement
2. Hostile belligerence
3. Paranoid projection
4. Grandiose expansiveness
5. Perceptual distortion
6. Anxious intropunitiveness
7. Retardation and apathy
8. Disorientation
9. Motor disturbances
10. Conceptual disorganization

Each type profile can be designated by a numerical label to indicate which mean syndrome scores are elevated 0.15 or more above the general mean. Thus a 1-10 profile means elevated "Excitement" (1) and "Conceptual disorganization" (10) scores. Similarly a 2-3 type profile implies that "Hostile belligerence" (2) and "Paranoid projection" (3) are elevated above the mean. This notation will be used hereafter as a shorthand but objective way to identify each type. Each subgroup has been labeled additionally by its most characteristic syndromes. For example the 2-3 subgroup members are called hostile-paranoids.

The Acute Psychotic Types

The patient sample, drawn from 16 State and university hospitals and clinics, con-

sisted of 374 men and 448 women. All patients manifested a functional psychosis, were hospitalized either for the first (54 percent) or second (41 percent) time, and were between 18 and 55 years of age. Cases with central nervous system or other neurological disorders, as well as alcoholics, were excluded. Interviews were conducted within 10 days of admission while patients were either not on tranquilizers or on a minimum drug dosage schedule. Approximately 75 percent were diagnosed schizophrenics, 23 percent as depressives, and 5 percent as manics. Of the total, 82 percent were white and the remainder of other ethnic origins.

A. Types among men

The mean syndrome score profiles for each of the types are given in tables 1 and 2. The labels applied are tentative and designed to be descriptive mainly of the salient syndromes. The male patient classes are as follows:

Excited (1-10) Type

The members of this group are characterized mainly by elevated excitement scores. While high scores on conceptual disorganization are frequent not all are above the mean. Occasionally patients in the group also manifest above-average scores on motor disturbances and grandiosity. All other syndromes average well below the general mean. Twenty-two percent of the excited type are diagnosed as manic.

Excited-Hostile (1-2-3-10) Type

The scores of all members are above the scale mean with respect to excitement and hostility. Elevated scores on paranoid projection are also common. Conceptual disorganization, although slightly elevated, is only slightly discriminating. Roughly two-thirds of the type were diagnosed as paranoid. It is thus likely that the class represents one variety of paranoid disturbance.

Hostile-Paranoid (2-3) Type

Most patients in this class have scores elevated on both hostile belligerence and paranoid projection. However, some members exhibit elevated scores mainly on hostile belligerence or mainly on paranoid projection. Excitement scores tend to be conspicuously low. Approximately 64 percent of the type are diagnosed paranoid type.

Hallucinated Paranoid (3-5-8) Type

The third paranoid class has as members patients scoring well up on paranoid projection and perceptual distortion. In addition some members score high on disorientation. This means that in addition to delusional misinterpretation of the action of others as persecutory or conspiratory, the type members also hear voices that accuse, threaten, or order. Diagnostically 71 percent of the group are categorized as paranoid type.

Grandiose-Paranoid (3-4-5) Type

All type members have extremely high scores on grandiose expansiveness. Most members, but not all, have in addition elevated scores on paranoid projection and perceptual distortion. In brief, members exhibit attitudes of self-importance and superiority, and report the possession of unusual gifts and powers. At times they identify with well-known personalities or claim special divine missions. Approximately 78 percent are diagnosed paranoid type.

Anxious-Depressed (6) Type

The members of this class all manifest above-average anxious intropunitiveness. A few patients also receive mildly elevated scores on retardation, perceptual distortion, or hostile belligerence. However, the intro-punitive score is always highest. Fifty-five percent of the group is diagnosed as depressed (psychotic or involutional) while the remainder are classed as schizo-affective,

acute undifferentiated, or paranoid. The intro-punitives are the most frequent of the types identified.

Retarded with Motor Disturbances

(7-9) Type

Members of this class all have scores elevated on retardation and apathy. In brief, they are seen as slowed in speech and movement; they may also whisper, block, or fail to answer at all. Many of the class members also manifest high motor disturbances scores. This means they may posture, manifest bizarre or manneristic movements, talk to themselves, and show muscular tension. The diagnoses for this group varied widely, the most common being acute undifferentiated (37 percent).

Disoriented (7-8-9-10) Type

This patient type is relatively small. Its members all have extreme disorientation scores. In addition some members have high scores on retardation, motor disturbances, and conceptual disorganization. The most frequent psychiatric diagnosis is schizophrenia, simple type (36 percent).

Anxious-Disorganized (3-5-6-7-8-9-10) Type

The most striking feature of this class is the presence of anxiety in conjunction with behaviors indicative of behavior disorganization or disintegration. All members of the type show elevated scores on anxious intro-punitiveness, perceptual distortion, and retardation. Some members also receive high scores on paranoid projection, disorientation, motor disturbances, and conceptual disorganization. Patient members are typically diagnosed as paranoid or as acute undifferentiated.

Thus the acute male psychotic sample can be viewed as including four kinds of patient classes. There is one excited type and four paranoid types (excited-hostile, hostile, grandiose, and hallucinated). The intro-punitiveness represent the anxious-depressive disorders. The behaviorally dis-

organized types are the anxious-disorganized who may represent transitional states, the disoriented, and the retarded-motor disturbed.

B. Types among women

Seven of the nine types isolated in the sample of women are essentially the same as those found among men. The excited (1-9-10) type closely resembles the male patient type except for a more elevated score on motor disturbances. As can be seen in table 2, the four female paranoid classes differ little from the corresponding groups identified in the male sample. There is an excited-hostile (1-2-3-10), a hostile-paranoid (2-3), an hallucinated paranoid (2-3-5), and a grandiose paranoid (3-4-5). The hostile paranoid among women is somewhat more paranoid than the male variety. The hallucinated paranoid more frequently manifests hostile behavior, while the grandiose type is less prone to exhibit excitement than corresponding male type members. Most of the paranoid types are diagnosed as paranoid.

The anxious-depressives (6) manifest a syndrome profile that is almost a duplicate of the corresponding male variety. The anxious-disorganized (3-5-6-7-9) resembles the correlative male class very closely with regard to paranoid projection, perceptual distortion, intro-punitiveness, and retardation: The female type, on the other hand, is distinctly less disoriented and less motor disturbed.

The female types that appear to be different from those found among men may now be described.

Excited-Disorganized (1-2-3-4-5-9-10) Type

All members of this class have elevated scores on excitement, paranoid projection, grandiosity, and conceptual disorganization. In addition, many score high on hostile belligerence, perceptual distortion (hallucinations), and motor disturbances. However, none of the group are disoriented.

Evidently the excited-disorganized represented an acutely disturbed group. Members are most likely to be diagnosed as manic or as schizophrenic, acute undifferentiated type. Further research is needed to determine whether the class is a transitional one.

Retarded-Disorganized (7-8-9) Type

Among men the disoriented (7-8-9-10) patient type was differentiated from the retarded-motor disturbed (7-9). Such a separation was not possible among women although both profiles appear among them. All members of the retarded-disorganized type have strongly elevated scores on retardation and apathy. In addition about two-thirds of the group also manifest high scores on disorientation and motor disturbances. Members tend to be diagnosed either as depressed or as catatonic.

The Chronic Psychotic Types

The second study was based on two groups of relatively chronic male psychotics with several previous hospital admissions. One sample of 207 was specially selected to assure representation of all likely sources of symptom variation. Another sample consisted of 359 schizophrenics newly admitted or readmitted to 32 Veterans Administration hospitals. In all, 44 State and federal institutions contributed to the sample. The modal patient was 36 years old, hospitalized three or four times previously for durations of 18 to 24 months. The age range was from 18 to 55. At the time of interview and rating patients were receiving at most either light sedatives or mild dosages of tranquilizer drugs.

Seven male patient types emerged from the analyses of the 3 chronic patient subsamples of 150. Their mean syndrome profiles are presented in table 3. Six of the chronic patient subgroups are substantially the same as corresponding acute patient subgroups. The retarded type represents the only new subgroup evolved. On the

other hand, the excited type and the hallucinated paranoid type failed to appear within the chronic subsamples. The extent of similarity between corresponding acute and chronic types as measured by the congruency coefficient is shown in table 5.

What differences are there in corresponding type profiles? The chronic excited-hostile subgroup is more often grandiose but less often paranoid than the acute subgroup. The chronic grandiose paranoid manifests slightly higher scores on excitement and hostile belligerence than the corresponding acute type. The corresponding intropunitive subgroups are quite similar in profile. The chronic anxious-disorganized type is less often disoriented than its correlative among acute patients. The chronic retarded-disorganized subgroup profile resembles the acute disoriented type ($C=0.86$). However, there are differences that preclude a match. Retardation scores are always elevated above the mean, while disorientation and motor disturbances may or may not be elevated in the retarded-disorganized subgroup. Among the acute disoriented type members, it is disorientation that is always elevated while retardation and motor disturbances are not. The closest match for the chronic retarded-disorganized type is the acute female type of the same name.

The retarded subgroup members exhibit high scores on retardation only. All other syndrome scores are typically below the mean.

The Long-Term Psychotic Types

While the so-called chronic sample included a high proportion of patients who had been hospitalized three and four times, emphasis was not on long-term hospitalization. In view of the high proportion of chronic long-term patients in psychiatric hospitals, further investigation of existing subgroups seemed to be of value. Accordingly a sample of 450 cases was drawn from data collected in connection

with 2 nationwide Veterans Administration studies of tranquilizer effects.

One-third of the 450 cases came from a drug discontinuation study (project 9). The average patient was 40 years old with 10 years of hospitalization. The most common diagnosis was schizophrenic reaction, paranoid type. Patients were interviewed either after 16 weeks on placebo or at the time of their clinical relapse following withdrawal of medication.

The remaining 300 cases were drawn from another VA drug study (project 14) of schizophrenics. The average patient was 41 years old with approximately 9.5 years of hospitalization. The most common diagnoses were chronic undifferentiated, paranoid, and hebephrenic. Each patient was interviewed and rated after 1 month on placebo capsules.

Seven subgroups were isolated among the long-term patients' subsamples. Their mean profiles are presented in table 4. Five of the types had been found in the acute and/or chronic samples, but the excited-disorganized and the hostile-motor disturbed were new. All members of the excited-disorganized subgroup have highly elevated scores on motor disturbances and conceptual disorganization. Nearly all have above average scores additionally on excitement. At the same time roughly two-thirds of the group score above average on retardation and apathy. The pattern is analogous to a catatonic excitement.

The hostile-motor disturbed subgroup is relatively small and defined mainly by elevated scores on hostile belligerence. Scores on motor disturbances are frequently, but not always, above the sample mean. Further data are needed to evaluate the stability of this type.

The similarity between five of the long-term subgroups and the acute and chronic subgroups can be assessed by examining table 5. The size of the congruency coefficient indicates rather strong agreement in the mean syndrome profiles.

An Acute Schizophrenic Sample

A study of newly admitted acutely disturbed schizophrenics was reported by McNair, Lorr, and Hemingway (1964). The data came from a phenothiazine study conducted by the Psychopharmacology Service Center of NIMH. A subsample of 150 men and 2 subsamples of 108 women were randomly selected for analysis. The age range was from 16 to 40 and the average age was 28. A clustering process was applied by hand to the two female subsamples. First a multiple cutting-score procedure was used to classify the male profiles. The procedure was followed by a search for new patient types.

The mean syndrome profiles for the types established are shown in table 6. The degree of similarity between the schizophrenic types and comparable acute psychotic types was evaluated by means of the congruency coefficient. (See table 7.) As may be seen five of the types agree very closely. There were several conditions that could account for the lack of agreement in the remaining types identified. First, the scoring key applied to the NIMH schizophrenics differed slightly from that applied to all other rating data. In all subsequent studies item No. 15 (attitude of superiority) was dropped from the excitement syndrome. Item No. 30 (slovenly appearance) was eliminated from motor disturbances and added to retardation. Thus three of the syndrome scores have slightly different bases. Another reason for the differences is that the cluster method applied was not as rigorous as in other analyses. The cross-cluster correlations of type (1-2-3-4) and of retarded-intropunitive are too high; these subgroups should have been eliminated. If these differences are considered, then the extent of agreement between the types in the two studies compared is high. Further, none of the types isolated is new or different from those found in the acute psychotic sample or the chronic psychotic sample.

It is of considerable interest to note that a sample of carefully diagnosed schizo-

phrenics includes nearly all psychotic types found among a random sample of acute psychotics. This finding supports the need for a reappraisal of currently accepted diagnostic classes.

Summary of psychotic types isolated

A summary of the findings is given on Table 8. The table shows that 12 subgroups have been indentified at least twice in the 4 studies described. Four paranoid groups emerged: Excited, hostile, grandiose, and hallucinated. These results suggest that the category schizophrenic reaction, paranoid type, is divisible in several subcategories. The excited type, to be found only in acute nonchronic samples, resembles the classical manic. Five subgroups can perhaps be categorized as disorganized in behavior in contrast to the paranoid groups. These are the retarded-motor disturbed, the disoriented, the retarded disorganized, the anxious disorganized, and the excited disorganized. The first two types may represent a further differentiation of the type called retarded-disorganized. Finally the classical depressive disorders appear to be represented by the intro-punitives and the retarded. However, reports by Grinker et al (1961) suggest that a more detailed study of the intro-punitive subgroup with use of inner state reports could reveal additional depressed subgroups.

Reliability of the types

Nine acute psychotic male types were established through application of the clustering procedure to the combined rater syndrome scores. The subgroups were then enlarged by adding members to each type if they satisfied the bounds set by mutiple cutting scores. Klett and McNair (1965) then tested type reliability by determining the extent to which interviewer and observer separately classified the same patients into each type. First the syndrome scores of the 374 males were standardized separately for the interviewer and the observer. Next the profiles for the interviewer were ma-

chine sorted into nine types using the multiple cutting score boundaries. The process was repeated for the set of observer scores, and the resulting classifications compared.

The findings were that altogether the interviewer and observer agreed in their rating of 73 percent of the 374 males. They jointly assigned 38 percent of the males to the same type category. The two judges also agreed that another 34 percent of the males were unclassifiable. Most of the disagreement between judges consisted in one rater classifying the patient when he could not be classified by the other rater's scores. Rarely did the two raters assign the same patient to different types.

Another problem was to determine how closely the combined rater classification could be approximated by the scores of only one rater. The analyses indicated that roughly 85 percent of the typing decisions would have been the same whether based on one or two raters.

Validity of the psychotic types

The evidence presented thus far has been supportive of the replicability of the types in a variety of patient samples. Data concerning reliability of the acute psychotic types have also been encouraging since disagreement between raters was found to be minimal. However, the types will be of little interest unless and until some evidence can be presented of their validity. What are the antecedents of the psychotic types? Can the types be differentiated with respect to social history, genetic background, or premorbid personality structure? Do the types respond differentially to available psychiatric treatment modalities? Are they useful for predicting length of hospitalization, duration of illness, or treatment response? Can concurrent validity be established through correlations with other measures of behavior disturbance?

The only validity evidence now available is given in a report by Klett and Lorr (1965) on the acute psychotic types. Identification of the psychotic types has been too recent

to permit accumulation of much validity data. In the study of the acute psychotics, data were collected relative to age, ethnic origin, religion, highest grade completed, marital status, residence (rural versus urban), occupation, diagnosis, duration of hospitalization, type of treatment received, and outcome status. Because of the relatively small number of cases within each of the nine types no statistical tests were attempted. The types were compared with the base rates for the variable and with each other. Thus any findings reported provide only exploratory and suggestive evidence of differences in background, clinical condition and outcome.

Table 9 presents a summary of the characteristics that appeared to differentiate the acute male types. Since a detailed report is available only two types will be described to illustrate the findings. For example, the anxious-depressed type members are 92 percent white. They are likely to be married currently and more likely than most to come from urban areas. Their fathers tend to have upper class jobs. Intropunitives are very likely to be diagnosed as depressive reaction, and to be placed on open wards. They receive psychotherapy more often than any other subgroup and tend more than most to get an approved discharge 6 months after admission. The disoriented subgroup, in contrast, are made up primarily of the less educated, lower class, unmarried, rural patients. Frequently mute during the interview, type members are usually diagnosed schizophrenic reaction, simple type, catatonic type, or chronic undifferentiated type. They had been ill a long time before they were committed involuntarily to a hospital and tended to stay hospitalized at 6 months far more than any other type.

A sizable number of IMPS ratings of newly admitted and chronic schizophrenic men were available from a series of drug evaluations included in the VA Cooperative Studies in Psychiatry. Two projects were concerned with newly admitted,

acutely ill psychotics while three projects sampled chronic long-term patients. The IMPS syndrome scores of 1,610 patients from these 5 studies were standardized on the same metric as the acute psychotic males described earlier. A multiple cutting score program was then applied to all cases. Of the total, 41 percent were classifiable, including 36 percent of the newly admitted and 46 percent of the chronic patients. Table 10 shows that the retarded-motor disturbed and the disoriented subgroups are likely to be chronic. The anxious subgroups and the paranoid subgroups are more likely to be newly admitted. The excited type is about equally chronic or acute. It is also noteworthy that of the classified chronic sample 76 percent are disoriented and retarded-motor disturbed.

What proportion of cases are classified within the limits set for members of a type? To answer this question a system of scoring was developed based on multiple cutting scores (MCS) for each type. From the distributions of the 10 syndromes, a cutting score or limit was set for each and every differentiating syndrome (nondiscriminating syndromes were ignored). The classification rules established for each type were applied sequentially by computer to all cases in the sample. The results on the acute psychotic sample indicated that approximately 60 percent of men and women could be classified. Cases unassigned by MCS were then correlated with the mean syndrome profile of each type. This procedure resulted in the addition of another 20 percent. Thus these nonoptimal procedures show that 80 percent of cases can be classified. Use of multiple discriminant functions would, of course, classify an even greater proportion of cases.

REFERENCES

- CATTELL, R. B. "A Note on Correlation Clusters and Cluster Search Methods." *Psychometrika*, 1944, 9, 169-184.
- CRONBACH, L. J., and GLESER, GOLDINE C. "Assessing Similarity Between Profiles." *Psychological Bulletin*, 1953, 50, 456-473.

HEINCKE, F. "Naturgeschichte des Hering. Die Lokalformen und die Wanderungen des Hering in der Europaaeschen Meeren." *Abhandlungen des Deutschen Seefischerei-Vereins*, 1898, 2, 1-223.

KLETT, C. J., and LORR, M. "Validity of the Acute Psychotic Types." Ch. 6, Lorr, M. (editor). *Explorations in Typing Psychotics*. Oxford: Pergamon Press. In press. 1965.

KLETT, C. J., and MCNAIR, D. "Reliability of the Acute Psychotic Types." Ch. 5, Lorr, M. (editor). *Explorations in Typing Psychotics*. Oxford: Pergamon Press. In press. 1965.

LORR, M. (editor). "Explorations in Typing Psychotics." Oxford: Pergamon Press. In press. 1965.

LORR, M., KLETT, C. J., and MCNAIR, D. M. "Syndromes of Psychosis." New York: the Macmillan Co., 1963.

LORR, M., KLETT, C., MCNAIR, D. M., and LASKY, J. J. "Inpatient Multidimensional Scale." Manual. Palo Alto: Consulting Psychologists Press, 1962.

NUNNALLY, J. "The Analysis of Profile Data." *Psychological Bulletin*, 1962, 59, 311-319.

McQUITT, L. L. "Pattern Analysis Illustrated in Classifying Patients and Normals." *Educational and Psychological Measurement*, 1954, 14, 598-604.

McQUITT, L. L. "Capabilities and Improvements of Linkage Analysis as a Clustering Method." *Educational and Psychological Measurement*, 1964, 24, 441-456.

SAUNDERS, D. S., and SCHUCHMAN, H. "Syndrome Analysis: An Efficient Procedure for Isolating Meaningful Subgroups in a Nonrandom Sample of a Population." Paper read at annual Psychonomic Society meeting. St. Louis, Mo., September 1962.

SAWREY, W. L., KELLER, L., and CONGER, J. J. "An Objective Method of Grouping Profiles by Distance Functions and Its Relation to Factor Analysis." *Educational and Psychological Measurement*, 1960, 20, 651-673.

STEPHENSON, W. "Some Observations on Q Technique." *Psychological Bulletin*, 1952, 49, 483-498.

THORNDIKE, R. L. "Who Belongs in the Family?" *Psychometrika*, 1953, 18, 267-271.

ZUBIN, J. "Socio-Biological Types and Methods for Their Isolation." *Psychiatry*, 1938, 237-247.

ZUBIN, J., FLEISS, J., and BURDOCK, E. L. "A Method for Fractionating a Population Into Homogeneous Subgroups." Unpublished paper, 1963.

TABLE 1.—Mean Syndrome Scores of 9 Acute Psychotic Types Among Male Patients

Type	Syndromes									
	EXC	HOS	PAR	GRN	PCP	INP	RTD	DIS	MTR	CNP
Excited.....	1.29	-0.65	-0.80	-0.32	-0.53	-0.68	-0.48	-0.28	-0.03	0.47
Excited-hostile.....	1.42	1.39	0.33	0.05	-0.51	-0.75	-0.74	-0.46	-0.29	0.34
Hostile paranoid.....	-0.52	0.80	0.17	-0.42	-0.55	-0.43	-0.55	-0.24	-0.40	-0.61
Hallucinated paranoid	-0.53	0.04	1.24	-0.23	1.32	-0.05	-0.24	0.22	-0.52	-0.12
Grandiose paranoid...	0.12	-0.41	0.20	2.02	0.51	-0.46	-0.45	-0.14	-0.61	-0.13
Anxious-depressed....	-0.46	-0.63	-0.80	-0.54	-0.43	1.15	-0.45	-0.45	-0.44	-0.62
Retarded-motor										
distrubed.....	-0.70	-0.83	-0.68	-0.56	-0.55	-0.30	1.43	0.06	0.52	-0.26
Disoriented.....	-0.43	-0.84	-0.93	-0.60	-0.38	-0.88	0.87	3.29	0.39	0.18
Anxious-disorganized	-0.50	0.10	1.28	-0.44	2.20	1.19	1.61	0.62	1.18	0.48

TABLE 2.—Mean Syndrome Scores of 9 Acute Psychotic Types Among Female Patients

Type	Syndromes									
	EXC	HOS	PAR	GRN	PCP	INP	RTD	DIS	MTR	CNP
Excited.....	1.49	-0.62	-0.89	-0.25	-0.55	-0.58	-0.70	-0.39	0.21	0.17
Excited-hostile.....	1.32	1.39	0.49	-0.04	-0.57	-0.49	-0.65	-0.39	-0.28	0.48
Excited-disorganized	2.13	0.54	1.40	2.28	1.16	-0.04	0.13	-0.20	1.51	1.81
Hostile paranoid.....	-0.63	0.70	0.41	-0.42	-0.53	-0.58	-0.59	-0.27	-0.43	-0.55
Hallucinated paranoid	-0.51	0.59	1.56	0.04	1.73	-0.44	-0.42	-0.24	-0.52	-0.56
Grandiose paranoid...	-0.15	-0.07	0.33	2.56	0.36	-0.79	-0.47	-0.05	-0.46	0.00
Anxious-depressed....	-0.46	-0.51	-0.67	-0.52	-0.45	1.13	-0.33	-0.35	-0.47	-0.63
Anxious-disorganized	-0.62	0.13	0.73	-0.38	1.74	1.20	1.23	-0.30	0.38	-0.33
Retarded-										
disorganized.....	-0.73	-0.92	-0.83	-0.51	-0.41	-0.25	1.50	1.13	0.38	-0.19

TABLE 3.—Mean Syndrome Scores of 7 Chronic Psychotic Types

Type	Syndromes									
	EXC	HOS	PAR	GRN	PCP	INP	RTD	DIS	MTR	CNP
Excited-hostile.....	1.88	1.08	0.02	0.19	−0.60	−0.46	−0.68	−0.39	−0.03	0.37
Grandiose paranoid...	0.77	0.35	0.70	2.51	0.79	−0.41	−0.46	−0.33	−0.02	0.66
Hostile paranoid.....	−0.32	1.28	0.94	−0.24	0.00	0.06	−0.33	−0.29	−0.45	−0.55
Anxious-depressed....	−0.51	−0.53	−0.74	−0.55	−0.34	1.54	−0.03	−0.34	−0.49	−0.66
Retarded.....	−0.74	−0.69	−0.84	−0.58	−0.46	−0.50	1.15	−0.38	−0.38	−0.45
Anxious-disorganized	−0.57	−0.04	1.03	0.06	2.02	1.70	1.03	−0.38	0.31	0.72
Retarded-disorganized.....	−0.41	−0.52	−0.60	−0.45	−0.15	−0.55	1.59	1.89	1.25	0.79

TABLE 4.—Mean Syndrome Scores of 7 Chronic Long-Term Psychotic Types

Type	Syndromes									
	EXC	HOS	PAR	GRN	PCP	INP	RTD	DIS	MTR	CNP
Excited-motor disturbed.....	1.27	−0.38	−0.36	−0.10	−0.34	−0.71	0.28	0.03	2.03	1.12
Hostile-motor disturbed.....	−0.26	0.66	−0.49	−0.46	−0.49	−0.36	−0.27	−0.46	0.33	−0.65
Hallucinated paranoid	−0.36	−0.44	0.71	−0.25	1.93	0.13	−0.46	−0.33	0.53	−0.59
Anxious-depressed....	−0.37	−0.87	−0.91	−0.35	−0.27	0.95	−0.47	−0.46	−0.44	−0.73
Retarded.....	−0.75	−0.98	−1.15	−0.52	−0.62	−0.68	0.21	−0.46	−0.94	−0.70
Retarded-motor disturbed.....	−0.53	−0.93	−1.01	−0.58	−0.61	−0.64	1.08	−0.37	0.86	−0.17
Disoriented.....	−0.48	−0.94	−0.99	−0.55	−0.59	−0.88	1.12	4.38	1.07	−0.11

TABLE 5.—Relations Between Profiles of Corresponding Psychotic Types in 3 Samples

Type	Patient samples		
	Acute versus chronic	Acute versus long term	Chronic versus long term
Excited-hostile.....	0.95		
Hostile paranoid.....	.75		
Grandiose paranoid.....	0.85		
Hallucinated paranoid.....		0.74	
Anxious-depressed.....	0.96	0.98	0.91
Anxious-disorganized.....	0.89		
Retarded-motor disturbed		0.93	
Disoriented.....		0.99	
Excited.....			
Retarded.....			0.89

TABLE 6.—Mean Standard Scores of 9 NIMH Patient Types

Type	Syndromes									
	EXC	HOS	PAR	GRN	PCP	INP	RTD	DIS	MTR	CNP
Excited-grandiose	1.4	0.0	0.0	2.0	0.0	−0.7	−0.7	−0.2	0.0	0.2
Excited-hostile	1.8	0.7	−0.3	−0.3	−0.7	−0.7	−0.9	−0.4	−0.1	−0.2
EXC-HOS- GRN-PAR	1.4	1.4	1.1	2.0	0.0	−0.9	−0.9	−0.3	−0.1	0.6
Hostile paranoid	0.0	1.2	0.9	−0.4	−0.4	−0.5	−0.7	−0.3	−0.4	−0.2
Hallucinated paranoid	−0.6	−0.2	1.2	−0.4	1.2	0.3	−0.6	0.0	−0.4	−0.6
Anxious-depressed	−0.2	−0.1	−0.7	−0.4	−0.6	1.2	−0.5	−0.3	−0.4	−0.4
Retarded- intropunitive	−0.8	−0.8	−0.4	−0.6	−0.5	0.7	0.4	−0.3	−0.7	−0.4
Retarded	−0.6	−0.2	−0.8	−0.5	−0.4	−0.5	1.2	−0.1	0.0	−0.2
Retarded- disorganized	0.0	−0.4	−0.8	−0.4	−0.4	−0.6	0.7	3.6	1.3	0.1

TABLE 7.—Relations Between Profiles of NIMH Schizophrenics and Acute Psychotic Types

Acute psychotic sample	C	NIMH schizophrenic sample
Excited-hostile	0.89	Excited-hostile.
Hostile paranoid	0.82	Hostile paranoid.
Hallucinated paranoid	0.92	Hallucinated paranoid.
Anxious-depressed	0.94	Anxious-depressed.
Disoriented	0.95	Retarded-disorganized.
Excited	0.40	Excited grandiose.
Grandoise paranoid	0.60	EXC-HOS-GRN-PAR.

TABLE 8.—Summary of Psychotic Types Isolated in 4 Studies

Type	Samples			
	Acute random	Chronic random	Long-term schizophrenics	Acute schizophrenics
Excited	(*)			(?)
Excited-hostile	(*)	(*)		(*)
Hostile paranoid	(*)	(*)		(*)
Grandiose paranoid	(*)	(*)		
Hallucinated paranoid	(*)		(*)	(*)
Anxious-depressed	(*)	(*)	(*)	(*)
Retarded-motor disturbed	(*)		(*)	
Disoriented	(*)		(*)	(*)
Retarded-disorganized	¹ ()	(*)		
Anxious disorganized	(*)	(*)		
Retarded		(*)	(*)	(*)
Excited-disorganized	¹ ()			(?)
Excited-motor disturbed			(*)	

¹ NOTE—Found only among women.

TABLE 9.—Summary of Characteristics of the Male Types

Type	Excited	Excited-hostile	Hostile-paranoid	Hallucinated-paranoid	Grandiose-paranoid	Anxious-depressed	Retarded-motor	Disoriented	Anxious-disorganized
Ethnic origin	White	White	City	1/3 Negro	1/3 Negro	White	Never married	Never married	
Religion	Protestant	Protestant	City	8th grade	Protestant	City	Rural	Rural	
Education	High school	Protestant	City	Most married	College or 8th grade	Upper class	Lower class	Few college	
Marital status	Most remarried	Protestant	City	Rural	Lower and middle class	Middle and upper class	Lower class	Never married	
Residence	Protestant	Protestant	City	Lower and middle class	Paranoid	Depressed	Acute undifferentiated	Simple	Paranoid and acute undifferentiated
Patient occupation	Protestant	Protestant	City	Paranoid	Paranoid	Depressed	Acute undifferentiated	Simple	Paranoid and acute undifferentiated
Father occupation	Protestant	Protestant	City	Paranoid	Paranoid	Depressed	Acute undifferentiated	Simple	Paranoid and acute undifferentiated
Diagnosis	Manic and paranoid	Paranoid	Paranoid	Paranoid	Paranoid	Depressed	Acute undifferentiated	Simple	Paranoid and acute undifferentiated
Commitment	Manic and paranoid	Paranoid	Paranoid	Paranoid	Paranoid	Depressed	Acute undifferentiated	Simple	Paranoid and acute undifferentiated
Ward type	Manic and paranoid	Paranoid	Paranoid	Paranoid	Paranoid	Depressed	Acute undifferentiated	Simple	Paranoid and acute undifferentiated
Duration of episode	> 10 months	Paranoid	Paranoid	Paranoid	Paranoid	Depressed	Acute undifferentiated	Simple	Paranoid and acute undifferentiated
Primary treatment	> 10 months	Paranoid	Paranoid	Paranoid	Paranoid	Depressed	Acute undifferentiated	Simple	Paranoid and acute undifferentiated
Adjunct treatment	Drugs	Paranoid	Paranoid	Paranoid	Paranoid	Depressed	Acute undifferentiated	Simple	Paranoid and acute undifferentiated
Outcome status	Drugs	Paranoid	Paranoid	Paranoid	Paranoid	Depressed	Acute undifferentiated	Simple	Paranoid and acute undifferentiated
Length of stay	Drugs	Paranoid	Paranoid	Paranoid	Paranoid	Depressed	Acute undifferentiated	Simple	Paranoid and acute undifferentiated
Age	(older)	Paranoid	Paranoid	Paranoid	Paranoid	Depressed	Acute undifferentiated	Simple	Paranoid and acute undifferentiated

TABLE 10.—Type Membership of Other Samples

Types (males)	Newly admitted		Chronic	
	N	Percent	N	Percent
Excited.....	32	52	29	48
Excited-hostile.....	28	76	9	24
Hostile paranoid.....	40	68	19	32
Hallucinated paranoid.....	16	70	7	30
Grandiose paranoid.....	27	79	7	21
Anxious-depressed.....	48	80	12	20
Retarded-motor disturbed.....	66	38	109	62
Disoriented.....	18	10	171	90
Anxious disorganized.....	15	83	3	17
Total.....	290	44	366	56

Supplementary Comments by Dr. Lorr
as Presented at the Conference

Our studies grew out of the frequently expressed dissatisfaction with the traditional psychiatric classificatory scheme, such as Kreitman and Beck have documented; the lack of reliability of specific classes. Others have provided ample evidence that the system is lacking in validity for treatment, outcome, or duration. Further, we were impressed by the fact that at no time has the system been rigorously tested or verified.

To quote Cameron: "It is important for persons working in the abnormal field to realize that the current official psychiatric classifications are not based on final and convincing scientific evidence. They are children of practical necessities. Decisions as to the group in which a given behavior disorder can fall depend upon schemata that actually were adopted, both in this country and in Great Britain, by a majority vote

of the practicing members of large associations."

In brief, the validity of the scheme is assessed in terms of the extent of agreement between the system and the clinical experience and conception of the clinical user. Certainly this extension of majority rule to scientific problems is a novel approach, to say the least.

Now, the initial goal set in our studies was to identify the major dimensions or syndromes descriptive of the domain of behavior disorders called psychotic. Our objective was not to replace the existing categories but to test them and to develop a useful research tool. The method of factor analysis was selected for several reasons. First, the concept of a syndrome is a statistical idea based on covariation. The view of the syndrome as an all-or-none disease entity we regard as an error and a carryover from physical medicine.

Secondly, a factorial experiment that yields a factor much resembles the operation of a clinician as he goes about detecting a syndrome. As the clinician interviews successive patients, he notes the simultaneous presence or absence of certain symptoms and behaviors, and eventually gives the functional unity observed a name. He does, in effect, an armchair factor analysis.

There are certain differences, however, in the two approaches. The clinician based his judgment on relatively unsystematic qualitative observations and on small and usually accidental samples of unspecified populations. The factorial experiment, on the other hand, is carried out on more carefully specified samples. Judgments are objective, fitted, qualified, and tested for reliability. Furthermore, all members of a sample are measured on all behaviors.

Finally, a computer assesses the degree of covariation and similarity among symptoms and groups those covarying together. The factorial procedure is thus far more easily verifiable and reproducible.

Now, the interview was chosen as the observational situation because it is here that the psychiatrist usually obtains the data for a diagnosis. However, recording was restricted to observable behavior and patient self-reports. Social history was ignored because it was deemed undesirable to sneak in etiology, onset, course, or chronicity by the back door. Such data only bias and confound the very outcomes the experiment is designed to achieve.

The scales constructed were designed to cover all relevant information concerning the psychotic disorders. The literature, including the APA Manual, was scrutinized as source material.

No less than five independent analyses have been made on sizable samples, on both male and female patients, within State, federal, and university settings. In general, 10 or 11 syndromes have been confirmed and reconfirmed. Because many of the syndromes are compelling and appear

congruent with psychiatric groupings, some critics argue that nothing has been contributed. This is clearly untrue. Some syndromes have been supported, others invalidated, and still other new categories have been established.

We also wish to emphasize that the descriptive scheme can be and has been rigorously checked. Now, factor analysis yields dimensions or symptom groupings. Emphasis is on the covariation among symptoms.

A second approach is to focus attention on groups of individuals who manifest similar or identical syndrome patterns. This process is designed to yield subgroups or types if they exist. Here factor analysis is unsuitable. Regardless, whether symptoms or people are factored, the operation still yields dimensions. Thus other statistical procedures, commonly called clustering techniques, must be applied instead. These are methods for sorting people into mutually exclusive categories.

Now, our report, the one that was handed out, represents a summary of findings of four studies. The general objective was to identify or recover existing subgroups. When one considers the psychotic population there seems to be ample justification for assuming that it is nonrandom; that it represents a mixture of several populations. Further, we sought to establish general-purpose categories that could be related backward to social history, onset, genetic background, and other antecedents and forward to treatment choices, duration, and the like.

Special-purpose treatment based categories are, in our view, likely to be ephemeral and trivial. General-purpose groupings are natural and are far more likely to lead to an understanding of the process and development of a disorder.

The 12 subgroups presented in the report were each found in at least 2 independent subsamples of 150 cases. A half dozen types were identified in virtually every analysis.

While one clustering procedure we devised was used in these studies, two independent methods have since been tested on the same data. Both Saunders' method and another technique of our own emerge with essentially the same groups. In an effort to see the effect of similarity indices we completed comparable analyses using person-person correlations, congruency coefficients, and distances between profiles. While there are slight differences the bulk of the findings were unchanged.

The question may properly be asked whether our types have been validated. The answer is "No," but we do have very interesting and suggestive supporting data. Since our analyses were completed only last June, there simply has not been time to validate these types as well.

Studies of Quantitative Approaches to Psychiatric Classification¹

John E. Overall, Ph. D.² and

Leo E. Hollister, M.D.³

In this paper we will review a program of research in which we have participated jointly during the last 6 years. A major interest throughout this period has been in exploring the utility of objective multivariate statistical and profile analysis methods for classification of psychiatric patients. We do not consider the results obtained so

far to be, in any sense, a final solution to the total complex problem of patient classification. The objectives of this research were never that presumptuous. Instead, we have concentrated from time to time on identifying and studying one or another major aspect of the total classification problem.

As we see it, any objective approach to patient classification involves three major aspects: (1) measurement, (2) nosology or typology, and (3) decision model. These we consider to be the principal ways in which one objective approach to patient classification can differ from another. Within this broad framework, a very large number of possible measurements, nosology, decision model combinations exist. We believe that a substantial effort will be required on the part of many researchers to evaluate the relative merits of alternatives available to us in each of these areas.

Before becoming engrossed in the details of special approaches, a brief outline of the several alternatives that we have considered may be in order. This outline is offered in the hope that our efforts will not be conceived as related to one particular system for patient classification.

1. Measurement:
 - (a) IMPS or MSRPP molecular symptom and behavior ratings.
 - (b) 16-dimensional BPRS.
 - (c) Four syndrome factors.
2. Nosology or typology:
 - (a) Standard nomenclature.
 - (b) Empirically derived modal clusters.
 - (c) Empirically derived drug response types.
3. Decision rules:
 - (a) Pattern probability or Bayesian conditional probability model.
 - (b) Multiple discriminant function D² model.

¹ This work was supported in part by Grant MYP-5144, Psychopharmacology Service Center, NIMH, and computer facilities were made available through Special Research Resources Grant USPHS 1PO 7-00024-01, NIH.

² Director, Research Computer Center, University of Texas Medical Branch, Galveston, Texas.

³ Veterans Administration Hospital, Palo Alto, Calif.

(c) Normalized vector product profile analysis.

(d) Profile correlation.

A particular quantitative approach to patient classification may include any combination of these measurements, patient typology, and decision-rule components. Our research endeavors have been concerned with constructing computer programs to classify patients on the basis of pretreatment rating profiles using various combinations of measurement, typology, and decision rules, and then with attempting to evaluate the relative merits of the different approaches. The response to different drug treatments has been used as a pragmatic criterion. In a very real sense, patients who responded differently to different drug treatments can be considered to be different kinds of patients. Furthermore, the practical implication of being able to classify patients into groups that respond especially well to one or another form of treatment makes the criterion of differential drug response an important one.

Measurement

As a basis for patient classification, the appropriate selection of measurement dimensions is of paramount importance. Psychiatric patients, like all humans, differ from one another in a myriad of ways. Most of these are irrelevant for problems of classification. A first step in definition of the measurement space involves an informed guess at the dimensions which are psychiatrically important in distinguishing one patient from another and in distinguishing the psychiatric patient from the normal person. Associated with this problem is consideration of how best to measure these psychiatrically important dimensions.

As we all know, there are several potentially useful approaches to psychiatric measurement. Partly out of coincidence, and partly because of the empirically demonstrated usefulness of psychiatric ratings as

a superior basis for evaluating treatment differences in controlled double-blind studies, we have chosen to work almost entirely with these types of measures. We believe that the rating scale is the most direct way to objectify and quantify the kinds of judgments upon which psychiatric decisions have traditionally been based.

The first two years of our endeavors, in close collaboration with Donald R. Gorham, can be characterized in retrospect as an attempt to understand the basic structure of the symptom measurement space (Gorham and Overall, 1960, 1961; Overall, Gorham and Shawver, 1961; Overall, 1962) (1, 2, 3, 4). What parsimonious set of dimensions is required in order to represent most of the objectively measurable differences among hospitalized psychiatric patients? How can these fundamental dimensions best be characterized in concepts meaningful to the trained psychiatrist or clinical psychologist? How can we best and most directly measure these fundamental patient differences? During this period, we used correlational analysis techniques, such as cluster analysis and factor analysis, along with a liberal interplay of clinical interpretation. The result was the Brief Psychiatric Rating Scale (BPRS) (Overall and Gorham, 1962) (5).

The BPRS contains sixteen 7-point rating scales for evaluating level of severity in 16 relatively independent symptom areas. The attempt has been to include a single scale to represent each factor dimension identified in previous analyses. Psychopathology can be characterized at various levels of abstraction. We conceive of a hierarchy ranging from molecular symptom and behavior items to global concepts of psychopathology. In this hierarchy of increasing abstraction, the BPRS rating constructs would seem to occupy an intermediate level. For purposes of patient classification enough detail is required to permit representation of important individual differences; yet enough parsimony must be retained to en-

able one easily to conceptualize the classification prototypes. We have found that the BPRS provides an adequate framework within which to view essential individual differences among hospitalized psychotic patients with which our work has been concerned. A facsimile of this scale is presented in figure 1. (*Note:* All figures and tables referred to herein may be found in section following references at the end of this paper.)

Before turning to discussion of problems of nosology and decision making, one or two important aspects of the measurement problem need to be mentioned briefly. The decision rules which we use in patient classification can, in each case, be related to an underlying mathematical or geometric model. The decisions which are reached in patient classification are based upon relationships derived from the abstract model. These decisions have meaning in the real world of psychopathology only to the extent that a one-to-one relationship has been established between the properties of the model and characteristics of the real world that the model is supposed to represent. The Pythagorean distance function and various vector product or profile correlation indices have meaning in terms of the same rectangular geometric model (Nunnally, 1962) (6).

Distances and relationships in the abstract geometric space are based upon coordinate axes which are orthogonal and equal in scale units. The results obtained from use of the abstract model will have meaning for the real world of psychopathology only if attempt is made to insure some degree of correspondence between these properties of the abstract model and characteristics of the measurement world. In short, use of some of these models implies statistical independence and comparability of scale units among profile scores (Overall, 1964) (7). In addition to these requirements, a meaningful origin for measurement scales is required by some models, e.g., normalized vector product

model. As we shall see, some approaches involve mathematical transformations of the original variables to meet requirements of the geometric model, e.g., multiple discriminant analysis.

The above assumptions are difficult to satisfy completely. The BPRS is far from a perfect instrument and can be recommended primarily because it has proven useful. The 16 scales were selected because they were found to represent meaningful and relatively independent symptom dimensions (higher correlations approximately 0.40 to 0.50, with majority close to zero). Each symptom scale has the same defensible natural origin of "not present." While we would not attempt to defend the equal-interval properties of these ordered category scales, each symptom construct is rated on exactly the same type of equal-appearing interval scale. As seen in table 1 (see "Note" on this page regarding reference to tables and figures), the variances of ratings are highly comparable for several scales across a heterogeneous patient population.

Means and standard deviations based on a sample of 549 clinically schizophrenic patients from 8 hospitals are shown in table 1. Correlations are presented in table 2. (We are indebted to Dr. Jonathan O. Cole for making available the data used in computing these means and correlations.) Perhaps the worst problem from the point of view of utilizing symptom intensity measures in statistical analysis is the lack of normality in distributions. For each symptom variable, a portion of the patients will have "not present" ratings. This is indicated by the fact that the mean values usually lie less than two standard deviation units from the zero point of the scale.

A powered vector factor analysis of the intercorrelation matrix yielded four orthogonal higher order factors. We have now analyzed intercorrelations based upon BPRS ratings from at least eight independent samples and extremely consistent results have been obtained. The four higher order factors can be adequately characterized as

thinking disturbance, psychomotor disturbance, paranoid interpersonal disturbance, and depressive disturbance. The variables related to each of these major independent psychiatric functions are shown in table 3.

Nosology or Typology

The next question we faced was that of defining nosologic types in terms of the measurement dimensions identified during the preliminary phase. As previously mentioned, we have taken several approaches to this problem.

Standard diagnostic nomenclature

Recognizing that there is currently much dissatisfaction with the standard nomenclature, we first became interested in attempting to make more objective and operational classification of patients according to conventional diagnostic concepts. The poor reliability and lack of validity of clinical diagnosis in psychiatry may be due to one of several factors, it was initially considered.

The difficulty may be in lack of consensual understanding, among experts, of the characteristics distinguishing various types—the nomenclature includes vague, disorganized, verbal characterizations consisting of admixtures of etiology, history, and symptomatology. Secondly, the difficulty may involve inability to evaluate reliably the characteristics that are important for diagnostic classification—psychiatric characteristics are not always most obvious and available to direct observation. Finally, the difficulty may simply be the lack of any clear-cut specification of decision rules for combining multiple observations in reaching a diagnostic decision—the standard nomenclature is completely silent on this important point.

The standard nomenclature represents a collection of descriptions of patient types. In the absence of biological, genetic, or biochemical evidence to the contrary, it is perhaps most realistic to consider this

nosology as a conceptual typologic system existing in the collective mind of psychiatrists. In our approach to this problem, we have sought an objective, quantitative definition of the conventional diagnostic types in terms of BPRS measurement dimensions.

A total of 38 experts, specialists of high reputation in psychiatry (28) and clinical psychology (10), were asked to conceive of a typical patient belonging to each of 13 functional psychotic types and to provide a BPRS rating profile representing the levels of severity of symptom that they would expect to be present in each type. This was a study of the consensual validity of psychiatric diagnostic nomenclature in terms of objective symptom-rating profiles. The method of obtaining these conceptual prototype profiles has been described in detail, together with some results obtained from them (Overall and Gorham, 1963; Overall, 1963) (8, 9). These prototype profiles which have been used as a basis for patient classification are presented in table 4.

The collection of 38 expert judge prototype profiles for each of 13 functional psychotic types provided data defining these 13 psychiatric populations. The prototypes derived in this way have served as a basis for classification of patients using various decision rules (Overall and Hollister, 1964) (10). In general, all of the classification approaches result in comparison of the individual patient's profile with each of the diagnostic prototype profiles and the assigning of the patient to the diagnostic class to which his profile is most similar. The several approaches differ in the way that "profile similarity" is defined.

Empirically derived modal clusters

The second approach that we considered disregarded standard nomenclature entirely. It started with the question: What types of patient profiles occur with beyond chance frequency in the patient population? The number of possible patterns of profile composed of 16 variables, each having 7 pos-

sible levels, is 7^{16} —a very large number. Not all of these many profiles occur with equal frequency. A modal type is a profile type occurring with greater frequency than most other profiles.

From a collection of several hundred BPRS rating profiles, six samples of clinically schizophrenic patients were selected. The 40 profiles in each sample were intercorrelated across the 16 symptom rating variables (Q-type correlations). The powered vector method of factor analysis (Overall and Porterfield, 1963) (11) was used to identify in each matrix the clusters of profiles which correlated most highly.⁴ Some analyses yielded four factors and some yielded five. A single "cluster profile" was computed to represent the profiles loading highly on each of the Q-type profile factors from the six different analyses. A "higher order" profile correlation matrix was computed using the cluster profiles derived from the several initial analyses. This matrix was factored by the same procedure in an effort to identify from the intercorrelations those modal types that appeared in more than one of the initial analyses. A total of five modal profiles were identified in the analyses of profiles for clinically schizophrenic patients.

Next, 4 samples of 40 BPRS profiles for clinically depressive patients were intercorrelated and factor analyzed by the powered vector method. Two of the 40 x 40 profile correlation matrices yielded four factors and two yielded three. The cluster profiles, computed as the mean of profiles

having high loadings on each of the factors, were intercorrelated and factor analyzed to identify types appearing in more than one analysis. Three modal types resulted from these analyses of profiles for clinically depressed patients.

Modal profiles resulting from the correlational analyses of 10 samples of 40 BPRS rating profiles are presented in table 5.

We must admit that the empirically derived modal types are somewhat difficult to conceptualize. This is no doubt due, in large part, to the fact that we bring into the situation preconceived stereotypes. Certain interesting symptom relationships are apparent. Anxiety is a marker which appears in both schizophrenic and depressed patients, and in each major clinical group we recognize one modal type that stands out by lacking anxiety as a prominent symptom. Hallucinatory behavior is confined to a single distinctive modal type. Hostility appears as a prominent feature in one schizophrenic and one depressive modal type. Depressive mood as a symptom is prominent in all three depressed types and notably low in schizophrenic modal types. Such differences would seem to have implications with regard to differential treatment responses of patients classified according to these modal prototypes. In a later section, differences in responses of different depressive modal types to drug treatment will be discussed.

Drug-response types

The third approach which was taken to definition of patient populations involved use of the pragmatic criterion of drug response. Patients who respond with exceptional improvement to treatment with different drugs should be considered to be fundamentally different kinds of patients.

There are several possible approaches one might use to the identification of drug-response types (if such types exist). Only those patients who respond best to each different treatment can be selected as samples from the response-type populations. Then

⁴ The factoring of a Q-type matrix involving 40 profiles with only 16 tests will be objected to by some. The first defense for the procedure is that factor analysis can be employed in a completely atheoretical manner for the practical purpose of locating clusters of tests that correlate highly. No geometric or mathematical consideration beyond the fact that certain profiles tend to correlate more highly than others need be invoked. Secondly, it can be shown that for every R-type analysis there is a corresponding Q-type analysis, and that the results of one can be obtained directly from the other. This being true, one analysis would seem as appropriate as the other. The factor analysis model would seem to require assumption of independence of test variables in Q-analysis, just as it requires assumption of independence of persons in R-analysis. With the BPRS ratings, this assumption is reasonably satisfied.

multivariate discriminant function techniques can be used to develop procedures for assigning the patient, on the basis of pretreatment measures, to the drug-response type his profile most closely approximates.

Alternatively, the familiar multiple regression method can be used to predict response of patients to drugs. Patients can be evaluated prior to treatment, and then change during treatment can be recorded. If change during treatment is related to pretreatment symptom characteristics, multiple regression analysis can be used to develop equations from which response can be predicted. As it turns out, the results from a number of such analyses indicate that overall response to treatment is highly related to pre-treatment symptom characteristics. An analysis of variance for overall difference between regression equations for the different drugs indicated highly significant differences in the several regression equations. Similar results have been obtained by other investigators using a similar analysis with different data (Goldberg, Cole, et al., 1964) (12).

The correlation of each of the pretreatment ratings to change in total pathology during treatment provides a meaningful picture of the types of patients responding best to each drug. Patterns presented in table 6 are based upon correlations of pretreatment ratings with change in comparable samples of clinically schizophrenic patients. It can be noted that the patterns of significant relationships are almost non-overlapping, emphasizing the significant differences between functions useful in predicting response to different drugs.

A simple stencil can be used to match patterns of symptom rating with the predictive patterns shown in table 6. This emphasizes the fact that the computer may be necessary in research phases and may be unnecessary in later clinical applications. Some interesting operations research needs to be done in the area of human simulation of computers. How can we display multiple measurement data to yield visual pattern

matching which corresponds best to results obtained using quantitative models and computer technology?

Decision Rules

The decision rules which we have programmed for computer use have been described previously (Overall and Hollister, 1964) (10). Regardless of the measurement instruments employed and the nosologic classes considered, the decision problem is one of judging the similarity of an individual patient profile to the prototype profiles for different classes. Differences in the procedures result from differences in the way that "profile similarity" is defined.

Pattern probability model

The pattern probability model has most frequently been called "Bayesian conditional probability model" in medical literature (Warner, et al., 1961; Overall and Williams, 1963). We prefer to use the term "pattern probability" to emphasize what we consider to be the central feature for psychiatric classification—i.e., the simple multiplicative rule for probability of a pattern of independent events. This model is particularly appropriate for use with categorical data in which there exists reasonable independence between variables. For a more detailed discussion of use of the pattern probability procedure for psychiatric classification, see Overall and Gorham (1962) (8).

Elementary probability theory defines the probability of joint occurrence of two or more independent events (e.g., a pattern of observations) as the product of their separate probabilities.

$$P_{y|k} = P_{x_1|k} \cdot P_{x_2|k} \cdots P_{x_n|k},$$

where $P_{y|k}$ is the probability of occurrence of pattern y in group k , and where $P_{x_1|k} \cdots P_{x_n|k}$ are simple probabilities associated with the separate symptoms in the pattern.

The probability that an individual will belong to population k and at the same time

have symptom pattern y is the joint probability P_{yk} .

$$P_{yk} = P_{P_y|k} \cdot P_k$$

The probability that a randomly selected patient who does have symptom pattern y will at the same time belong to diagnostic population k is given by the ratio of the probability of being from group k and having pattern y to the sum of probabilities of having pattern y in all populations.

$$P_{k|y} = \frac{P_{yk}}{P_{y1} + P_{y2} + \cdots + P_{yk} + \cdots + P_{yn}}$$

If we can obtain estimates of the simple probabilities of occurrence of various separate symptom ratings in each of the several populations, we can use this model to estimate the probability with which any particular patient comes from each population. The assumption of independence among profile components within diagnostic populations must be recognized using this model. It will be noted that the assumption of independence applies to relationships within classes; thus, the correlation between variables across a heterogeneous population (table 1) are not necessarily in violation of this assumption. Partitioning of the total population into more homogeneous groups tends to reduce correlations among variables. Advantages of this model include appropriateness of its use with qualitative data involving no more than nominal scale properties.

Multiple discriminant function with D^2

The multiple discriminant function is theoretically the most general procedure making use of the Pythagorean distance model. A transformation of original scores into statistically independent "canonical variates" (each a weighted sum of the original profile scores) insures conformity to assumptions of the geometric model no matter what the variances and intercorrelations of the original scores so long as they can be assumed to be equal from one group to the next. In the multiple discriminant function

approach, the new weighted mathematical variables are computed in such a way that the average squared distance, D^2 , between diagnostic groups is maximized relative to variation within the groups. This tends to insure maximum separation of scores by individuals in different populations; thus, patients tend to be more easily identified as belonging to a particular population.

The weighting coefficients which yield new mathematical variables having maximally different values between groups are computed as solution of the matrix equation $(B - \lambda W)a = O$. A technical discussion of the method of analysis is presented by Rao (1952) (13). Results of the application of the method to a study of the configural relationships among psychiatric diagnostic stereotypes has been presented elsewhere (Overall, 1963) (9).

Once the multiple discriminant analysis procedure has been used to define new uncorrelated composite variables, these new variables can be taken as coordinate axes in a rectangular geometric space. Distances between profiles in this rectangular space can be computed using a generalization of the simple Pythagorean theorem—e.g., "the square on the hypotenuse is equal to the sum of the squares on the other two sides." More generally, the square of the distance between two profile points is equal to the sum of squares of differences in scores on uncorrelated, comparable unit profile components. More will be said about the geometric representation of profile vectors in the next section.

It is important to note that the multiple discriminant analysis, and Mahalanobis D^2 , employs a mathematical transformation of the original scores so that the (new mathematically derived) variables actually used in computing inter-profile distances have statistical properties identified with characteristics of the geometric model. It cannot be emphasized too strongly that the usefulness of abstract models depends upon bringing into some sort of relationship the

properties of the model and the characteristics of the real world that the model is supposed to represent. If such one-to-one relationships are not established to begin with, then deductions from the model will have no more implications for the real world than the characteristics of the real world had for the model to start with. The Mahalanobis D^2 approach requires fewer assumptions concerning characteristics of the profile scores because the calculations insure correspondence between the data and the model. Implicit is the assumption of some sort of meaningful interval scale of measurement (which may be different for each profile variable).

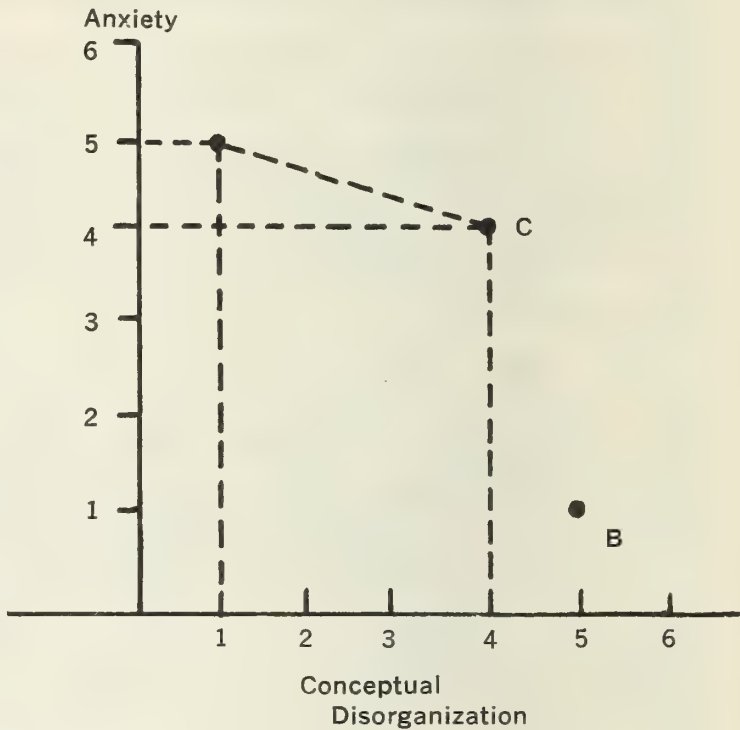
Normalized vector product model

Psychologists have for some time discussed an index of profile similarity which we will call the d^2 statistic (Osgood and Suci, 1952; Cronbach and Gleser, 1953) (14, 15). This simple index of profile similarity is based upon the same rectangular geometric model previously mentioned. It differs from the Mahalanobis D^2 by not correcting for correlations which may exist among profile components. Thus, if it can be assumed that the profile components are uncorrelated within groups and have equal variances, the simple d^2 statistic and the more involved Mahalanobis D^2 will turn out to be identical. The fact is that we can always assume independence and use the d^2 statistic. If the profile scores are not too highly correlated, good results can be expected. The d^2 statistic is mentioned, not because we have used it except on rare occasion, but because of its close relation to procedure which we found very useful.

If we conceive of profile scores as representing coordinate values in a multidimensional rectangular coordinate system, we can locate each rating profile as a point in that space. For example, in the simplest case involving only two-variable profiles, each individual can be located according to his scores.

	Anxiety	Conceptual disorganization
Profile A	5	1
Profile B	1	5
Profile C	4	4

The three two-component profiles A, B, and C can be located as points in a two-dimensional plane with coordinate axes corresponding to "anxiety" and "conceptual disorganization."



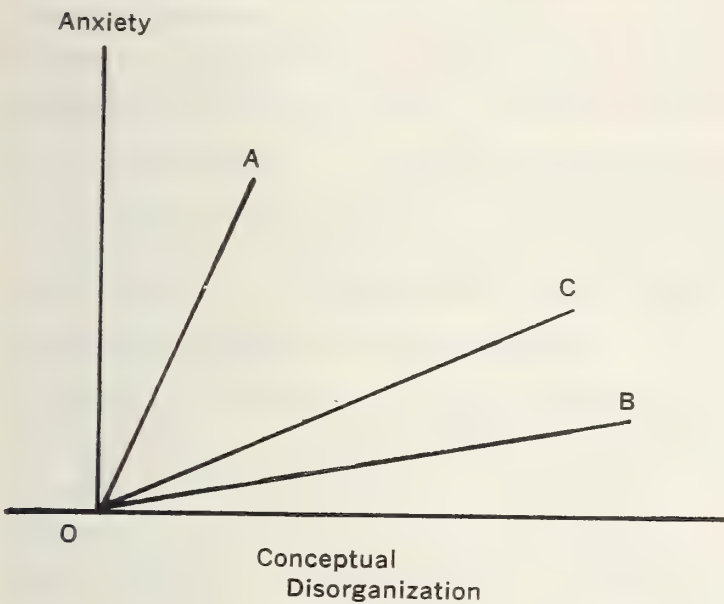
The distance between A and C can be computed from the sum of squares of differences in projections (scores) on the coordinate axes.

$$d^2 = d_1^2 + d_2^2 = (5 - 4)^2 + (4 - 1)^2 = 10$$

In a similar manner, the distances between all possible pairs of profiles can be computed. Although we cannot visualize the relationships directly in multidimensional space, we can appreciate that exactly the same kinds of operations can be used to evaluate distances between profiles involving numerous components.

There is another way that relationships between profiles can be conceived in this rectangular geometric space. Each profile can be presented as a vector radiating out from origin with end-point determined by projections on the coordinate axes, as before. The vector picture for the three

hypothetical two-component profiles considered above can be obtained simply by connecting the profile points with the origin.



Viewed in this way, we recognized another meaningful way to conceive of relationships between profiles—e.g., in terms of angular relationships among profile vectors. When we recall that a triangle can be completely defined in terms of two sides and the included angle, it is obvious that the distance AC contains no information not present in the lengths and angle between the profile vectors. It can be shown that the squared distance AC is equal to the sum of the square vector lengths minus twice the product of vector lengths multiplied by the cosine of the angle between the vectors. Because of correspondences established between the geometric model and properties of the profile measurements, we can write this relationship as follows (Nunnally, 1963; Overall, 1964).

$$d^2 = \sum a_i^2 + \sum c_i^2 - 2\sqrt{\sum a_i^2} \sqrt{\sum c_i^2} \cos \Theta$$

where a_i are scores in profile A and c_i are scores in profile C.

We see that the distance between profile end-points depends upon both the sums of squares and the sum of cross-products of profile elements. It is interesting to consider, not the distance between profile points (vector end-points), but simply the

differences in directions of profile vectors. This is tantamount to scaling all profile vectors to unit length and then examining the distances between them.

The normalized vector product index of profile similarity is simply the sum of cross-products of scores in two profiles after the scores in each have been scaled so that the sum of squares is unity. Computationally, this value can be obtained as follows:

$$\cos \Theta = \frac{\sum a_i c_i}{\sqrt{\sum a_i^2} \sqrt{\sum c_i^2}}$$

where a_i are scores in profile A and c_i are scores in profile C.

In terms of the normalized profile vectors, the squared distance between normalized vector end-points is given in the following simple equation.

$$d^2 = 2(1 - \cos \Theta)$$

The normalized vector product approach is conceptually equivalent to projecting all profile vectors to the surface of a unit-radius hypersphere and then examining the distances between the vector end-points. All of this is simply to say that there is a direct relationship between the vector product (and profile correlation) approach and the familiar Pythagorean geometric distance function. Because of this direct relationship, one cannot overlook the importance of assumptions of independence, comparable scale units and, in the case of the raw-score normalized vector product, meaningful zero points for measurement scales.

Profile correlation

The simple product moment correlation coefficient provides another frequently used index of profile similarity. The more similar the pattern of scores in two profiles, the higher will be the correlation between them.

No additional space will be devoted to discussion of this index because it can be regarded as a special case of the normalized vector product, and all geometric considerations associated with that model extend to profile correlations. The difference between

the normalized vector product and profile correlation is in the location of the origin. If each profile component is expressed as a deviation score about the profile mean, then the normalized vector product of the deviation scores is precisely the product moment correlation between the raw-score profiles.

The computational and conceptual similarities between the normalized vector products and profile correlation models should not be used to gloss over the very considerable practical differences between them. As Cronbach and Gleser (1953) (15) have described, the profile correlation eliminates average level or average severity as a factor in distinguishing between profiles. Two profiles are identical by this method if the pattern in one is identical to the pattern in the other. In the normalized vector product approach, two profiles are considered identical if all points in one are proportional to all points in the other. These two approaches, thus, result in quite different partitioning of the total measurement space. (Empirical evidence is beginning to suggest that the proportional equivalence is most compatible with clinical conception of psychopathology. As a patient improves or gets worse, his symptomatology changes proportionately, e.g., symptoms he did not have initially do not increase in absolute severity equal to those symptoms he had most to start with. The change tends to be proportionate, with most severe symptoms changing most.)

Applications in Drug Research

While we would not consider that a completely adequate system for psychiatric classification can be based solely upon patient responses to current therapeutic techniques, patients who respond differently to different forms of treatment are in a very real sense different types of patients. Different drugs act in different manners at different neurophysiological levels. Study of responses of patients to psychotropic drugs

should help us to retain meaningful distinctions present in the current nosology and can, perhaps, suggest modifications that are needed.

In the course of our studies of drug responsiveness, we have used standard nomenclature prototypes, empirically derived modal clusters, and empirically derived drug-response types. An examination of results obtained using each of these approaches will be included here. The drug studies were conducted as a cooperative project within eight Veterans Administration hospitals and were supported by grant MYP-5144, NIMH. Clinical investigators responsible for the project include Drs. J. L. Bennett, George Katz, Isham Kimbell, Jack Shelton, Gene Caffey, Veronica Pennington, Merlin Johnson, and Fred Meyer.

While the several drug treatment groups were not studied concurrently, the studies followed a similar general protocol with sampling from the same patient population. Patients in the studies of antipsychotic drugs can be considered to represent random samples from the population of newly admitted (or readmitted) male schizophrenic patients in the Veterans Administration. Antidepressant drugs have been studied in clinically depressed patients. The research design has routinely called for pre- and post-treatment ratings by two independent observers using the 16-variable Brief Psychiatric Rating Scale (BPRS).

Standard Nomenclature Prototypes

The first approach to patient-type specificity in drug action involved use of prototype profiles based upon standard diagnostic concepts. The development of these standard nomenclature prototypes has been described in an earlier section of this paper. Distance function analyses of the prototype profiles indicated three major syndrome clusters which can be adequately described as paranoid, schizophrenic, and depressive. Because these 3 major syndromes account for most of the differences in symptom con-

figurations of all 13 original subtypes and because small samples available in clinical drug research permit us to look only at gross differences, the three derived major syndrome prototypes rather than the original 13 subtypes have been used in studying drug responses. When we use the terms "paranoid," "schizophrenic," and "depressive" in this section, we are not referring to clinical diagnoses. We are explicitly talking about patients having 16-variable BPRS symptom profiles most like the three prototype profiles shown in figure 2.

Data from seven drug groups were used to evaluate the potential usefulness of the three major syndrome profile types. For each patient in each treatment group, pretreatment rating profiles were compared with the three major syndrome prototypes using the normalized vector product model. Each patient was assigned on the basis of his pretreatment rating profile to either paranoid, schizophrenic, or depressive, depending upon the largest normalized vector product similarity index.

Without going into detail with respect to all results, the index of total pathology will be selected as an example. Simple change scores were computed by subtracting pretreatment ratings from post-treatment ratings to obtain a single comprehensive measure of improvement during treatment. Mean improvement by major syndrome class is shown for seven drugs in table 7. In terms of "total pathology" improvement, there is no significant overall difference in improvement between profile types. There is a significant overall difference between drugs. Most important, significant drug \times patient-type interactions are observed among certain pairs of drugs. The phenothiazine derivatives plus triperidol appear relatively effective in patients with paranoid-type pretreatment rating profiles. Although the effectiveness of these particular drugs is rather stable across patient types, it appears generally least in the depressive subclass. In contrast, oxypertine and

imipramine have least effectiveness in patients with paranoid-type profiles and greatest effectiveness in clinically schizophrenic patients with depressive-type profiles. These results suggest the possible pragmatic value of this approach to patient classification. They also appear to support the validity of paranoid, schizophrenic, and depressive distinctions within psychiatric nosology.

Empirically Derived Modal Clusters

Experience in controlled double-blind studies of the so-called anti-depressant drugs has been largely disappointing. One after another supposedly effective anti-depressant has been found to be little or no better than placebo in samples of clinically depressed patients (Overall et al., 1962, 1964; Hollister et al., 1963, 1965) (16, 17, 18, 19). One possible reason for these poor results in the face of general clinical acceptance of these drugs may be that we are really treating a heterogeneous mixture of patients with diverse treatment needs.

In a double-blind controlled study imipramine and thioridazine were compared in groups of carefully classified depressed patients. Only those patients having BPRS profiles most like the standard nomenclature prototypes for depression than like any other subclass profiles were included in the depressive samples. The results failed to support the significant superiority of imipramine in this sample of carefully classified depressed patients, and in fact the observed mean improvement tended to favor thioridazine (Overall et al., 1964) (19).

Empirical profile analysis techniques were subsequently used to determine whether patients within the general depressive class tend to cluster in profile configurations. The results indicated three distinct homogeneous profile types within the general depressive class—anxious depression, hostile depression, and retarded depression.

As it turned out, most of our study patients were of the anxious type.

In table 8 are mean change scores representing improvement in total pathology for anxious, hostile, and retarded depressive types during treatment with thioridazine and imipramine. While there is no overall significant difference between the two drugs in the total sample of depressed patients, a significant drug \times patient-type interaction is present. Thioridazine appears to be the drug of choice for anxious depressives, and imipramine appears to be the drug of choice for retarded depressives. A closer examination of the data indicates that this result is not due simply to specific actions of the drugs on anxiety and motor retardation factors, although these are strong contributing factors in the total responses. This same interaction pattern is observed when the core symptom depressive mood is considered separately.

These results strongly suggest the practical need to distinguish between anxious-agitated depressive types and retarded depressive types. Although adjectives such as agitated, retarded, reactive, and endogenous are frequently employed with reference to depression, they are not recognized in the standard nomenclature. These findings, including relevance in drug research, suggest how empirical results may lead to modifications which need be made in the standard nomenclature.

Multiple Discriminant Analysis of Drug Response Types

Multiple discriminant analysis (Rao, 1952, p. 364) is a powerful statistical technique which can provide answers to three important questions: (a) Do patients who respond well to one drug differ from those who respond well to another? (b) How many different symptom types are needed to represent differences among specificities of several drugs previously studied? (c) What is the configuration of response types for the several drugs previously studied? For this

approach, drug response was used as the criterion for selecting patients. Response types identified in the analysis are based directly upon the results of drug studies.

For purposes of this analysis, only patients who evidenced more than 15 points improvement (single rater) in BPRS total pathology were selected for inclusion. In our experience, 15 points total pathology is about the average improvement associated with a good antipsychotic drug in our patient population. This means that these patients, selected on the basis of their drug responses, can be considered to represent random samples from the population of patients who do respond actively to each of several drugs. Seven groups of patients showing good response to one of seven different drugs were selected for the analysis.

The first question asked in the analysis was whether patients in samples selected because of high response to one of seven different drugs differed significantly in pretreatment rating profiles from one group to the next. An approximate overall test of significance can be used where one has large degrees of freedom within groups (Rao, 1952, p. 372). For the present analysis, the pooled within groups N was only 145, not really a large value. To compensate for this small N , one should require a smaller p -value than he otherwise would in order to remain conservative. As it turned out, a normal deviate z -score value of 9.5 was obtained. Even with small degrees of freedom, this value (which would be expected by chance less than one time in a million) is so extreme that we conclude that patients who responded well to the different drugs did differ in pretreatment rating profiles.

The second question that was asked concerned the number of different dimensions required to represent significant differences among the high-response patients in the several different drug groups. This question is related to the question of how many types are required to represent differences in patients responding differently to different

drugs. It may be that the same type responds well to several different drugs, but that another type responds to other drugs. In answer to this question, three dimensions of profile differences were found, indicating separation of four drug-response types.

The final problem involves study of configural relations among patients responding well to the different drugs. It was found that patients selected as responding well to acetophenazine and perphenazine did not differ significantly in pretreatment profile characteristics. That the same type of patient responds well to two similar phenothiazines is not a disconcerting result. It was found in this analysis also that patients responding well to oxypertine did not differ significantly in pretreatment symptom configuration from those responding well to perphenazine and acetophenazine. Oxypertine was withdrawn from the market because the manufacturer felt that, while it appeared to be an effective drug, it did not have enough uniqueness to recommend its addition to the already long list of antipsychotic drugs. Thus, high-response patients to three drugs did not differ significantly and form the nucleus for one distinct drug-response type. Differences in these results and the patterns displayed in table 7 are not due solely to the different methods of analysis, but are also due in part to the fact that in the present analysis only the high-response patients were considered.

Patients responding well to benzquinamide differed significantly in pretreatment profiles from those responding well to acetophenazine, perphenazine, or oxypertine. This drug, representing a distinctive chemical class, may well have different locus of action. Not significantly different from the benzquinamide responders were patients who responded well to thioridazine. While thioridazine is a phenothiazine, it has a piperazine ring which has recognized potential for modifying its action. The benzquinamide-thioridazine re-

sponders form a second discrete profile type.

Patients responding well to trifluoperidol differed from patients responding to each of the other compounds. Patients responding to imipramine differed from patients responding to each of the other compounds. Thus, each of these drugs resulted in definition of a distinctive drug-response type.

The configuration of similarities and differences between pre-treatment characteristics of patients responding well to the seven drugs are presented in figure 3. The inset of orthogonal vectors I, II, and III represents the approximate directions of the three significant discriminant functions in the three-dimensional space. The first discriminant function contrasts imipramine and trifluoperidol with the other drugs. The second discriminant function contrasts benzquinamide and thioridazine with the other drugs. The third discriminant function contrasts imipramine with all of the other drugs.

We feel that the results of this analysis are among the most meaningful that we have obtained. Four different types of patients were identified, each type responding best to a different class of drugs. The configuration of drugs, in terms of behavior specificity, is satisfyingly in agreement with pharmacologic and chemical differences. We believe that this approach has much to recommend it as a basis for further work in psychiatric typology using drug response as a technique for identifying patient populations.

Summary and Conclusions

Statistical evidence indicates relationship between pretreatment symptom characteristics and response to treatment. The type of patient that responds best to one drug appears different from the type that responds well to another drug. If stable functions having the same degree of discrimination as these initial findings can be established, the practical value could

amount to an improvement of as much as 50 percent in overall therapeutic efficacy of psychiatric drug treatment. The results certainly suggest that a serious error might be made by premature conclusion that classification of psychiatric patients is not worthwhile.

REFERENCES

(1) GORHAM, D. R., and OVERALL, J. E. "Drug Action Profiles Based Upon an Abbreviated Psychiatric Rating Scale." *J. Nerv. Ment. Dis.*, 1960, 132, 528-535.

(2) GORHAM, D. R., and OVERALL, J. E. "Dimensions of Change in Psychiatric Symptomatology." *Dis. Nerv. System*, 1961, 22, 576-580.

(3) OVERALL, J. E., GORHAM, D. R., and SHAWVER, J. R. "Basic Dimensions of Change in Symptomatology of Chronic Schizophrenics." *J. Abnorm. Soc. Psychol.*, 1961, 62, 597-602.

(4) OVERALL, J. E. "Dimensions of Manifest Depression." *Psychiat. Res.*, 1962, 1, 239-245.

(5) OVERALL, J. E., and GORHAM, D. R. "The Brief Psychiatric Rating Scale." *Psychol. Rep.*, 1962, 10, 799-812.

(6) NUNNALLY, J. "The Analysis of Profile Data." *Psychol. Bull.*, 1962, 59, 311-319.

(7) OVERALL, J. E., "Note on Multivariate Methods for Profile Analysis." *Psychol. Bull.*, 1964, 61, 195-198.

(8) OVERALL, J. E., and GORHAM, D. R. "Pattern Probability Model for Classification of Psychiatric Patients." *Behav. Science*, 1963, 8, 108-116.

(9) OVERALL, J. E. "Configural Analysis of Psychiatric Diagnostic Stereotypes." *Behav. Science*, 1963, 8, 211-219.

(10) OVERALL, J. E., and HOLLISTER, L. E. "Computer Procedures for Psychiatric Classification." *J. Am. Med. Asso.*, 1964, 187, 583-588.

(11) OVERALL, J. E., and PORTERFIELD, J. L. "Powered Vector Method of Factor Analysis." *Psychometrika*, 1963.

(12) GOLDBERG, S. C., COLE, J. O., and KLERMAN, G. L. "Differential Prediction of Improvement Under Three Phenothiazines," In J. R. WITTENBORN and P. A. MAY (editors) "Prediction of Individual Differences in Response to Pharmacotherapy." Chas. C. Thomas in press.

(13) RAO, C. R. "Advanced Statistical Methods in Biometric Research." New York: Wiley & Sons, 1952.

(14) OSGOOD, C. E., and SUCI, G. J. "A Measure of Relation Determined by Both Mean Differences and Profile Information." *Psychol. Bull.*, 1952, 49, 251-262.

(15) CRONBACH, L. J., and GLESER, G. C. "Assessing Similarities Between Profiles." *Psychol. Bull.*, 1953, 50, 456-473.

(16) OVERALL, J. E., and others. "Drug Therapy in Depressions. Controlled Evaluation of Imipramine, Isocarboxazide, Dextroamphetamine-amobarbital and Placebo." *Clin. Pharm. Therap.*, 1962, 3, 16-23.

(17) HOLLISTER, L. E., and others. "Evaluation of Desimipramine in Depressive States." *J. New Drugs*, 1963, 3, 161.

(18) HOLLISTER, L. E., and others. "Controlled Comparison of Amitryptiline, Imipramine and Placebo in Hospitalized Depressed Patients." *J. Nerv. Ment. Dis.*, 1965 (in press).

(19) OVERALL, J. E., and others. "Imipramine and Thioridazine in Depressed and Schizophrenic Patients." *J. Am. Med. Asso.*, 1964, 189, 605-610.

TABLE 1.—Means and Standard Deviations of BPRS Ratings for 549 Clinically Schizophrenic Patients in NIMH Cooperative Study

	Mean	SD
Somatic concern.....	2.2	1.7
Anxiety.....	3.4	1.9
Emotional withdrawal.....	3.6	2.0
Conceptual		
Disorganization.....	3.8	2.0
Guilt feelings.....	2.0	1.6
Tension.....	3.7	1.6
Mannerisms-posturing.....	2.2	1.7
Grandiosity.....	2.2	1.8
Depressive mood.....	2.5	1.6
Hostility.....	2.9	2.0
Suspiciousness.....	4.2	2.0
Hallucinatory behavior....	3.4	2.3
Motor retardation.....	2.2	1.7
Uncooperativeness.....	2.8	1.9
Unusual thought content..	4.4	2.1
Blunt affect.....	3.4	1.9

TABLE 2.—Intercorrelations Among 16 BPRS Variables Based Upon 549 Patients From NIMH Cooperative Study

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1.....	1.00															
2.....	0.21	1.00														
3.....	1.00		1.00													
4.....				1.00												
5.....					1.00											
6.....						1.00										
7.....							1.00									
8.....								1.00								
9.....									1.00							
10.....										1.00						
11.....											1.00					
12.....												1.00				
13.....													1.00			
14.....														1.00		
15.....															1.00	
16.....																1.00

TABLE 3.—Projections of 16 BPRS Symptom Ratings on 4 Orthogonal Factors From NIMH 8 Hospital Study

	I	II	III	IV
Somatic concern	0.10	−0.05	0.04	0.25
Anxiety	0.24	−0.11	0.22	0.71
Emotional withdrawal	0.02	0.90	0.00	0.02
Conceptual disorganization	0.46	0.21	0.22	−0.06
Guilt feelings	0.09	−0.09	0.04	0.73
Tension	0.12	0.03	0.33	0.46
Mannerisms-posturing	0.02	0.53	0.01	−0.01
Grandiosity	0.27	−0.11	0.32	−0.29
Depressive mood	0.06	0.05	−0.02	0.80
Hostility	−0.03	−0.06	0.88	−0.02
Suspiciousness	0.30	−0.13	0.79	0.04
Hallucinatory behavior	0.82	0.01	−0.16	−0.02
Motor retardation	−0.07	0.65	−0.29	0.05
Uncooperativeness	−0.13	0.57	0.45	0.09
Unusual thought content	0.82	−0.08	0.18	−0.04
Blunted affect	0.05	0.82	−0.05	0.00

TABLE 4.—*Prototype Profiles for 13 Functional Psychotic Classes Described in the Standard Nomenclature*

	1. Somatic concern	2. Anxiety	3. Emotional withdrawal	4. Conceptual disorganization	5. Guilt feelings	6. Tension	7. Mannerisms and posturing	8. Grandiosity	9. Depressive mood	10. Hostility	11. Suspiciousness	12. Hallucinatory behavior	13. Motor retardation	14. Uncooperativeness	15. Unusual thought content	16. Blunted affect
1. Paranoia.....	1.86	1.92	1.94	1.17	0.72	2.44	0.61	4.31	0.75	4.19	4.94	0.61	0.19	2.89	3.92	1.56
2. Paranoid state.....	2.18	3.00	2.13	2.71	1.39	3.29	.71	3.42	1.42	4.03	4.87	1.55	.42	3.47	3.97	1.61
3. Paranoid schizophrenic.....	3.32	3.16	3.24	3.49	1.30	3.57	1.43	4.30	1.41	4.84	5.49	3.62	.68	4.08	4.78	2.38
4. Acute undifferentiated schizophrenic.....	3.24	4.37	3.68	4.76	2.45	4.24	2.84	2.05	2.18	3.13	3.61	4.00	1.53	3.37	4.63	2.76
5. Catatonic schizophrenic.....	1.66	2.79	5.42	4.68	2.47	3.68	5.53	1.61	1.76	3.13	2.97	4.13	5.16	4.87	4.58	4.18
6. Hebephrenic schizophrenic.....	2.21	1.74	4.89	5.50	1.18	2.71	5.00	2.32	1.18	2.26	2.58	4.74	1.39	3.47	5.29	4.71
7. Simple schizophrenic.....	1.92	1.37	4.47	3.39	1.05	1.50	2.18	.84	1.13	1.39	1.97	1.50	2.03	2.32	2.58	4.61
8. Chronic undifferentiated schizophrenic.....	2.47	2.16	3.21	2.79	1.58	2.13	1.71	1.29	1.37	1.79	2.45	1.29	1.24	1.74	2.32	2.95
9. Residual schizophrenic.....	2.89	2.16	4.32	4.11	1.92	2.53	2.97	2.18	1.61	2.50	3.21	3.18	1.95	2.84	4.03	4.16
10. Schizo-affective.....	2.92	3.68	2.65	3.68	3.30	4.00	1.81	2.16	3.49	2.86	3.00	2.68	1.86	2.59	3.92	2.03
11. Psychotic depressive.....	4.53	4.79	3.34	3.29	5.00	3.82	1.95	.39	5.47	2.76	2.29	1.82	4.13	3.03	3.47	2.05
12. Manic-depressive, depressive.....	4.55	4.84	3.66	2.42	5.37	3.37	1.71	.45	5.53	2.00	2.21	1.21	5.03	2.74	2.87	1.63
13. Manic-depressive manic.....	.68	.63	1.61	3.82	.58	4.74	2.32	5.26	.42	3.76	2.13	1.16	0	3.50	2.79	.92

TABLE 5.—Modal Types Derived From Analysis of Similarities Among Patient Profiles

	Somatic concern	Anxiety	Emotional withdrawal	Conceptual disorganization	Guilt feelings	Tension	Mannerism-posturing	Grandiosity	Depressive mood	Hostility	Suspiciousness	Hallucinatory behavior	Motor retardation	Uncooperativeness	Unusual thought content	Blunted affect
Withdrawn-disorganized, schizo-type S1	0.8	0.8	3.0	4.5	0.3	1.2	1.8	1.7	0.4	0.4	1.6	2.2	1.0	0.7	3.6	3.8
Thinking disturbance, schizo-type S2	1.6	3.9	2.2	2.5	1.9	3.0	1.3	.3	2.4	1.8	4.5	4.3	1.2	1.1	3.9	2.3
Anxious-tense-disorganized, schizo-type S3	1.7	3.0	2.7	2.8	1.0	3.1	1.4	.5	2.0	1.2	1.6	.6	1.6	1.4	2.2	3.1
Withdrawn-tense-suspicious, schizo-type S4	1.4	2.6	3.0	2.5	1.2	3.1	1.6	1.3	1.3	2.8	3.2	1.1	.9	2.3	1.7	2.0
Anxious-tense-unusual thinking, schizo-type S5	1.8	3.9	1.3	2.2	2.7	3.5	1.0	1.8	2.0	2.1	2.7	.7	.6	.7	3.7	1.5
Anxious-tense depression, type D1	2.5	4.3	.4	.1	1.5	3.0	.2	0	4.5	.4	.3	0	1.8	.2	.1	1.4
Anxious-hostile depression, type D2	3.6	3.4	1.9	.4	.8	2.0	1.1	0	3.7	3.4	1.8	.3	1.3	1.1	.6	1.9
Withdrawn-retarded depression, type D3	1.4	1.9	4.2	2.4	2.8	1.9	0	0	4.1	.4	.3	.2	3.5	.9	.4	3.7

TABLE 6.—Correlations Between Pretreatment Symptom Ratings and Change in Total Pathology (7 Drug Groups)

	Acetophenazine	Perphenazine	Oxypertine	Perphenazine-Amitryptiline	Imipramine	Thioridazine	Trifluoperidol
Somatic concern.....	0.29	−0.08	0.16	0.20	0.25	0.34	0.23
Anxiety.....	0.20	0.16	0.42	0.25	0.37	0.34	0.03
Emotional withdrawal.....	0.15	0.44	0.36	0.16	0.18	0.37	0.30
Conceptual disorganization.....	0.33	0.43	0.10	0.33	0.04	0.21	−0.01
Guilt feelings.....	0.16	0.16	0.42	0.06	0.15	0.26	0.10
Tension.....	0.33	0.29	0.30	0.19	0.08	0.50	0.11
Mannerisms-posturing.....	0.26	0.47	0.26	0.24	0.09	0.49	0.15
Grandiosity.....	0.10	0.16	−0.42	0.03	−0.21	−0.04	0.42
Depressive mood.....	0.15	0.04	0.34	0.07	0.28	0.31	−0.02
Hostility.....	0.05	0.26	0.15	−0.04	0.02	0.24	0.24
Suspiciousness.....	0.37	0.27	0.13	0.13	0.22	0.10	0.51
Hallucinatory behavior.....	0.33	0.22	0.37	0.21	0.09	0	0.45
Motor retardation.....	0.22	0.24	0.46	0.27	0.49	0.02	0.16
Uncooperativeness.....	0.20	0.34	0.24	0.25	0.05	0.32	0.30
Unusual thought content.....	0.43	−0.03	0.30	0.20	−0.23	0.07	0.17
Blunt affect.....	0	0	0.11	−0.15	0.04	.03	0.26

TABLE 7.—Mean Improvement in BPRS Total Pathology for 7 Drug Treatments

	Para-noid	Schizo-phrenic	Depres-sive
Acetophenazine.....	30.8	28.5	24.8
Perphenazine.....	30.0	25.8	14.2
Oxypertine.....	13.5	29.5	35.7
Perphenazine-amitryptiline..	21.2	24.1	24.1
Imipramine.....	11.3	15.4	24.5
Thioridazine.....	24.3	22.4	25.9
Triperidol.....	46.3	23.9	21.2

TABLE 8.—Mean Changes in BPRS Total Pathology for 3 Depressive Subtypes

	Imipramine	Thioridazine
Anxious.....	7.0	21.5
Hostile.....	22.6	23.6
Retarded.....	19.0	2.2

BRIEF PSYCHIATRIC RATING SCALE

OVERALL AND GORHAM

DIRECTIONS: DRAW A CIRCLE AROUND THE TERM UNDER EACH SYMPTOM WHICH BEST DESCRIBES THE PATIENT'S PRESENT CONDITION.

1. SOMATIC CONCERN - DEGREE OF CONCERN OVER PRESENT BODILY HEALTH. RATE THE DEGREE TO WHICH PHYSICAL HEALTH IS PERCEIVED AS A PROBLEM BY THE PATIENT, WHETHER COMPLAINTS HAVE REALISTIC BASIS OR NOT.

NOT PRESENT VERY MILD MILD MODERATE MOD. SEVERE SEVERE EXTREMELY SEVERE

2. ANXIETY - WORRY, FEAR, OR OVER-CONCERN FOR PRESENT OR FUTURE. RATE SOLELY ON THE BASIS OF VERBAL REPORT OF PATIENT'S OWN SUBJECTIVE EXPERIENCES. DO NOT INFER ANXIETY FROM PHYSICAL SIGNS OR FROM NEUROTIC DEFENSE MECHANISMS.

NOT PRESENT VERY MILD MILD MODERATE MOD. SEVERE SEVERE EXTREMELY SEVERE

3. EMOTIONAL WITHDRAWAL - DEFICIENCY IN RELATING TO THE INTERVIEWER AND THE INTERVIEW SITUATION. RATE ONLY DEGREE TO WHICH THE PATIENT GIVES THE IMPRESSION OF FAILING TO BE IN EMOTIONAL CONTACT WITH OTHER PEOPLE IN THE INTERVIEW SITUATION.

NOT PRESENT VERY MILD MILD MODERATE MOD. SEVERE SEVERE EXTREMELY SEVERE

4. CONCEPTUAL DISORGANIZATION - DEGREE TO WHICH THE THOUGHT PROCESSES ARE CONFUSED, DISCONNECTED OR DISORGANIZED. RATE ON THE BASIS OF INTEGRATION OF THE VERBAL PRODUCTS OF THE PATIENT; DO NOT RATE ON THE BASIS OF THE PATIENT'S SUBJECTIVE IMPRESSION OF HIS OWN LEVEL OF FUNCTIONING.

NOT PRESENT VERY MILD MILD MODERATE MOD. SEVERE SEVERE EXTREMELY SEVERE

5. GUILT FEELINGS - OVER-CONCERN OR REMORSE FOR PAST BEHAVIOR. RATE ON THE BASIS OF THE PATIENT'S SUBJECTIVE EXPERIENCES OF GUILT AS EVIDENCED BY VERBAL REPORT WITH APPROPRIATE AFFECT; DO NOT INFER GUILT FEELINGS FROM DEPRESSION, ANXIETY, OR NEUROTIC DEFENSES.

NOT PRESENT VERY MILD MILD MODERATE MOD. SEVERE SEVERE EXTREMELY SEVERE

6. TENSION - PHYSICAL AND MOTOR MANIFESTATIONS OF TENSION, "NERVOUSNESS", AND HEIGHTENED ACTIVATION LEVEL. TENSION SHOULD BE RATED SOLELY ON THE BASIS OF PHYSICAL SIGNS AND MOTOR BEHAVIOR AND NOT ON THE BASIS OF SUBJECTIVE EXPERIENCES OF TENSION REPORTED BY THE PATIENT.

NOT PRESENT VERY MILD MILD MODERATE MOD. SEVERE SEVERE EXTREMELY SEVERE

7. MANNERISMS AND POSTURING - UNUSUAL AND UNNATURAL MOTOR BEHAVIOR, THE TYPE OF MOTOR BEHAVIOR WHICH CAUSES CERTAIN MENTAL PATIENTS TO STAND OUT IN A CROWD OF NORMAL PEOPLE. RATE ONLY ABNORMALITY OF MOVEMENTS; DO NOT RATE SIMPLE HEIGHTENED MOTOR ACTIVITY HERE.

NOT PRESENT VERY MILD MILD MODERATE MOD. SEVERE SEVERE EXTREMELY SEVERE

8. GRANDIOSITY - EXAGGERATED SELF-OPINION. CONVICTION OF UNUSUAL ABILITY OR POWERS. RATE ONLY ON THE BASIS OF PATIENTS STATEMENTS ABOUT HIMSELF OR SELF-IN-RELATION-TO-OTHERS, NOT ON THE BASIS OF HIS Demeanor IN THE INTERVIEW SITUATION.

NOT PRESENT VERY MILD MILD MODERATE MOD. SEVERE SEVERE EXTREMELY SEVERE

9. DEPRESSIVE MOOD - DESPONDENCY IN MOOD, SADNESS. RATE ONLY DEGREE OF DESPONDENCY; DO NOT RATE ON THE BASIS OF INFERENCES CONCERNING DEPRESSION BASED UPON GENERAL RETARDATION AND SOMATIC COMPLAINTS.

NOT PRESENT VERY MILD MILD MODERATE MOD. SEVERE SEVERE EXTREMELY SEVERE

10. HOSTILITY - ANIMOSITY, CONTEMPT, BELLIGERENCE, DISDAIN FOR OTHER PEOPLE OUTSIDE THE INTERVIEW SITUATION. RATE SOLELY ON THE BASIS OF THE VERBAL REPORT OF FEELINGS AND ACTIONS OF THE PATIENT TOWARD OTHERS; DO NOT INFER HOSTILITY FROM NEUROTIC DEFENSES, ANXIETY NOR SOMATIC COMPLAINTS. (RATE ATTITUDE TOWARD INTERVIEWER UNDER "UNCOOPERATIVENESS".)

NOT PRESENT VERY MILD MILD MODERATE MOD. SEVERE SEVERE EXTREMELY SEVERE

11. SUSPICIOUSNESS - BELIEF (DELUSIONAL OR OTHERWISE) THAT OTHERS HAVE NOW, OR HAVE HAD IN THE PAST, MALICIOUS OR DISCRIMINATORY INTENT TOWARD THE PATIENT. ON THE BASIS OF VERBAL REPORT, RATE ONLY THOSE SUSPICIONS WHICH ARE CURRENTLY HELD WHETHER THEY CONCERN PAST OR PRESENT CIRCUMSTANCES.

NOT PRESENT VERY MILD MILD MODERATE MOD. SEVERE SEVERE EXTREMELY SEVERE

12. HALLUCINATORY BEHAVIOR - PERCEPTIONS WITHOUT NORMAL EXTERNAL STIMULUS CORRESPONDENCE. RATE ONLY THOSE EXPERIENCES WHICH ARE REPORTED TO HAVE OCCURRED WITHIN THE LAST WEEK AND WHICH ARE DESCRIBED AS DISTINCTLY DIFFERENT FROM THE THOUGHT AND IMAGERY PROCESSES OF NORMAL PEOPLE.

NOT PRESENT VERY MILD MILD MODERATE MOD. SEVERE SEVERE EXTREMELY SEVERE

13. MOTOR RETARDATION - REDUCTION IN ENERGY LEVEL EVIDENCED IN SLOWED MOVEMENTS AND SPEECH, REDUCED BODY TONE, DECREASED NUMBER OF MOVEMENTS. RATE ON THE BASIS OF OBSERVED BEHAVIOR OF THE PATIENT ONLY; DO NOT RATE ON BASIS OF PATIENT'S SUBJECTIVE IMPRESSION OF OWN ENERGY LEVEL.

NOT PRESENT VERY MILD MILD MODERATE MOD. SEVERE SEVERE EXTREMELY SEVERE

14. UNCOOPERATIVENESS - EVIDENCES OF RESISTANCE, UNFRIENDLINESS, RESENTMENT, AND LACK OF READINESS TO COOPERATE WITH THE INTERVIEWER. RATE ONLY ON THE BASIS OF THE PATIENT'S ATTITUDE AND RESPONSES TO THE INTERVIEWER AND THE INTERVIEW SITUATION; DO NOT RATE ON BASIS OF REPORTED RESENTMENT OR UNCOOPERATIVENESS OUTSIDE THE INTERVIEW SITUATION.

NOT PRESENT VERY MILD MILD MODERATE MOD. SEVERE SEVERE EXTREMELY SEVERE

15. UNUSUAL THOUGHT CONTENT - UNUSUAL, ODD, STRANGE, OR BIZARRE THOUGHT CONTENT. RATE HERE THE DEGREE OF UNUSUALNESS, NOT THE DEGREE OF DISORGANIZATION OF THOUGHT PROCESSES.

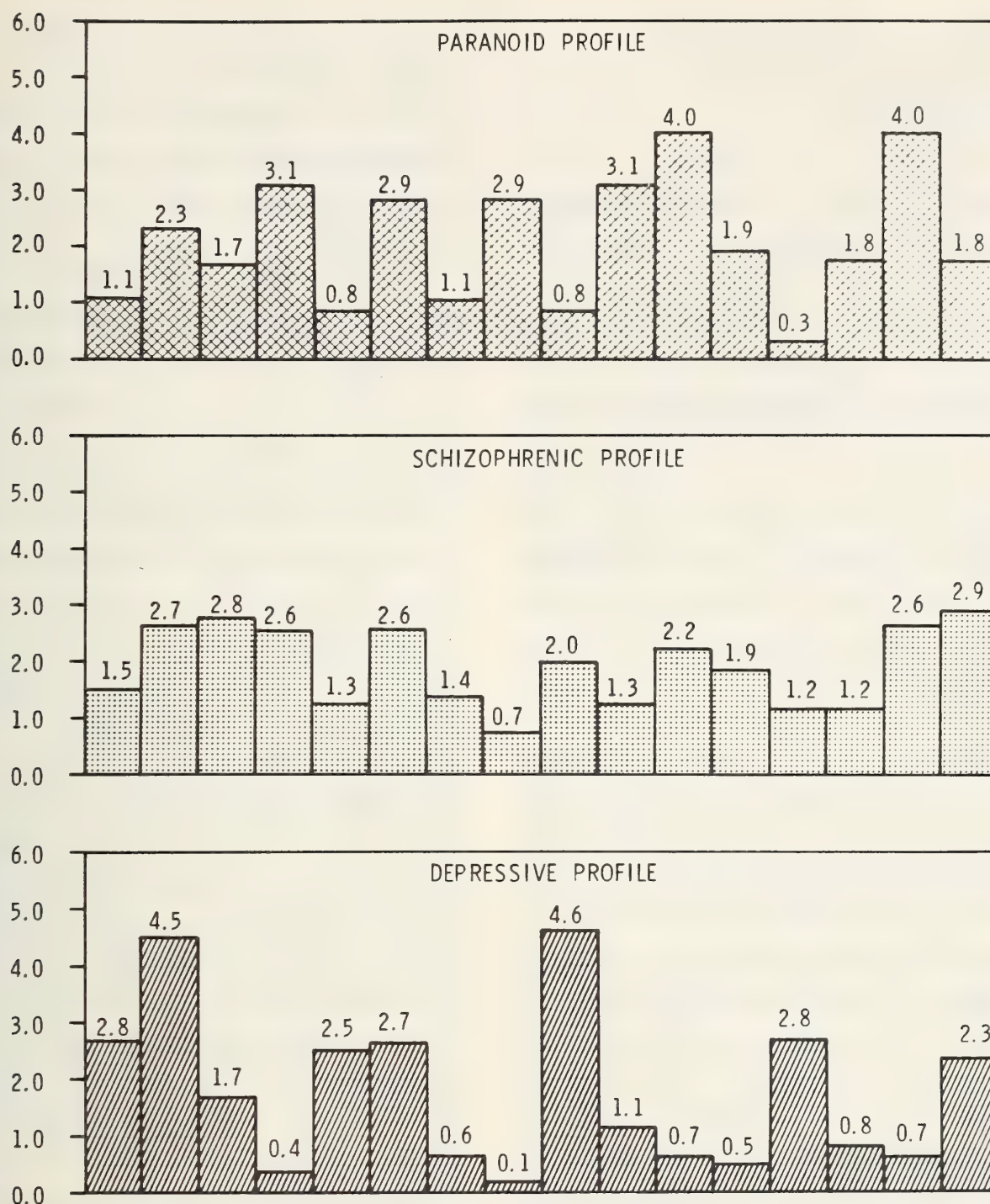
NOT PRESENT VERY MILD MILD MODERATE MOD. SEVERE SEVERE EXTREMELY SEVERE

16. BLUNTED AFFECT - REDUCED EMOTIONAL TONE, APPARENT LACK OF NORMAL FEELING OR INVOLVEMENT.

NOT PRESENT VERY MILD MILD MODERATE MOD. SEVERE SEVERE EXTREMELY SEVERE

© J. E. OVERALL & D. R. GORHAM

FIGURE 1.—Brief psychiatric rating scale (Overall and Gorham).



The 16 bars in each profile represent 16 BPRS variables in the order listed in Figure 1.

FIGURE 2.—Prototype profiles for paranoid, schizophrenic, and depressive classes.

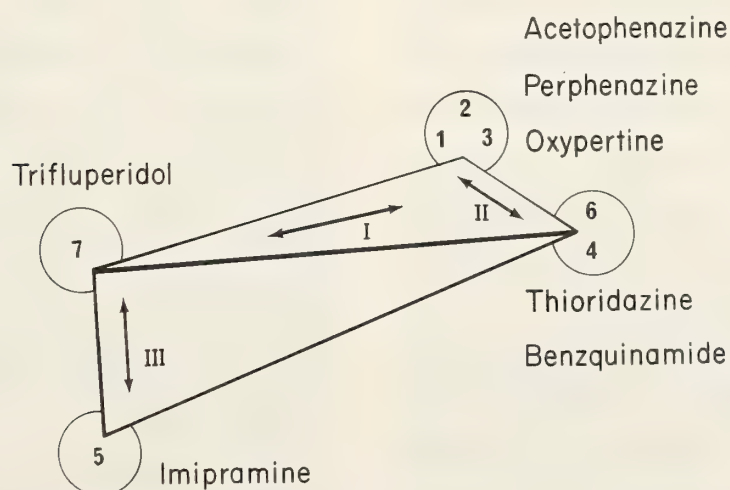


FIGURE 3.—Heuristic model derived from multiple discriminant analysis of relationships among rating profiles of patients who responded well to seven different drugs.

Supplementary Comments by Dr. Overall as Presented at the Conference

I would like to say that this conference has been very stimulating and important to us in providing an impetus for us to sit down and attempt to integrate within a single conceptual framework the results from a variety of different analytic approaches that we have employed in the past.

Our paper has summarized the results of a number of these different analyses. It is not always easy to bring into consistent relationship the results of such varied approaches, and I must admit that our paper was aimed more at describing different kinds of results than at attempting to distill from them a few simple guidelines which may be of some immediate usefulness in considering possible revisions in concepts of psychiatric nosology.

Since our paper has presented a majority of the empirical backgrounds for what I have to say, I would like to take this opportunity to summarize what we now consider to be the most important implications of this research for problems of classification in psychiatry. In this discussion I will follow Dr. Torgerson's excellent distinction between dimensional and typologic or structural descriptions of psychopathology.

From a dimensional point of view, a very consistent pattern of relationships can be observed to exist among psychiatric symptoms. We tend to think of these as syndromes, since they represent clusters of symptoms which tend to appear together. The four major syndromes can be described as depressive disturbance, paranoid interpersonal disturbance, thinking disturbance, and primary withdrawal-retardation.

I call your attention to table 3 in our paper as an example of one of many anal-

yses from which these four primary dimensional constructs were derived.

If one considers these to be the primary structural dimensions of the rating space, it becomes meaningful to classify patients according to the relative prominence of these four major syndrome characteristics. This would lead to four major classification regions within the psychiatric symptom space:

- A. Depression
- B. Paranoid interpersonal disturbance
- C. Withdrawal-retardation syndrome
- D. Thinking disturbance syndrome

But now from a typologic point of view, our research has led us to recognize that actually occurring patient types do not correspond to this simple breakdown. Other analyses have shown us that clinical diagnostic concepts cut across these primary dimensions. For example, in repeated typologic analyses of similarities and differences among clinically depressed patients, we have consistently found what might be considered to be mixed types of depression; e.g., a simple depression with anxiety, depression with paranoid interpersonal disturbance, and depression with withdrawal-retardation.

I call your attention to the last three profiles at the bottom of table 5 as indications of the results of these typological analyses of rating profiles from clinically depressed patients.

Our results have indicated that it makes considerable difference for the treatment of psychiatric depressions just what the concomitant symptom complex is. For example, anxious depressions respond well to thioridazine while retarded depressions respond poorly. On the other hand, anxious depressions respond poorly to imipramine or amitryptiline but retarded depressions respond well.

The problem of describing complexities in subclass types can be handled adequately, it would appear, by considering both a primary and secondary syndrome. For ex-

ample, the four major structural types degenerate into combinations of the following types: A pure depression or simple depressive type, but also depression with paranoid interpersonal disturbance and depression with withdrawal and retardation. A simple paranoid interpersonal disturbance syndrome, but also a paranoid syndrome with thinking disturbance and a paranoid syndrome with depression.

In addition to these four major syndrome types and the combinations observed in profile cluster analyses, we must add a fifth primary syndrome and two combinations. The additional syndrome has to do with agitation-excitement such as might be associated, in purest form, in the classic manic reaction. Thus, we would add to the list of symptom constellations a pure agitation-excitement, agitation-excitement with depression, and agitation-excitement with thinking disturbance.

The symptom measurement space has four primary structural dimensions which we can describe as an agitation versus retardation dimension, a depression versus elation dimension, a paranoid hostility versus compliance dimension, and a thinking disturbance dimension. Since elation, compliance, and extreme rational ego-controlled thought processes do not tend to get the individual into social difficulties, we tend not to consider these among the major psychopathologies. Nonetheless, they

do exist in the four-dimensional measurement space. They may be much more important in evaluation of neurotic disorders than in characterization of so-called psychotic disorders. At any rate we still have a simple four-dimensional structural model.

Within the simple four-dimensional phenomenological model, we recognize some 12 to 15 regions which appear important for patient classification. The importance of these regions is predicted on the results of typological analyses of profile similarities and upon the recognition that patients in these different regions tend to respond differently to different drug treatments.

It will be recognized that the system of nosology that we have proposed is based entirely upon phenomenological differences among psychiatric patients. Using the profile assignment method discussed in our paper for this conference, it is possible to specify precisely rules for assigning patients among these various subclasses for research purposes, since, in fact, the subclasses have been identified and defined in terms of psychiatric rating profiles. Since psychiatric ratings merely represent explicit recording of clinical psychiatric observations, it seems obvious that classification into these subgroups can be accomplished clinically by direction of attention to the four primary dimensions of phenomenological difference among patients having major psychiatric illness.

A Phenomenological Typology of Schizophrenia¹

*Martin M. Katz, Ph. D.*²

In an earlier paper, a rationale was presented for the need to develop new systems of patient classification which would be more objective than the present psychiatric diagnostic scheme, and would have more promise in research (Katz and Cole, 1963). In that report we described the development of a classification of acute schizophrenics which was based on ratings by significant relatives of the symptomatic and social behavior of these patients just prior to their hospitalization. The typology was phenomenological in nature; it was based on the observations of lay people. The system was derived from information having to do with the manifest behavior of the patient only, and it ignored the other aspects of the clinical history and course of illness, which are considered to be of prime diagnostic significance. In this paper, I will report on the developments of this research since that initial investigation. To provide some perspective on the types of studies which have evolved, I will indicate some of the thinking which determined the directions that the research has taken.

Differing theoretical and methodological approaches have resulted over the years in the development of several typologies in the fields of personality and the mental disorders. These systems are defended primarily on the basis of theory, their coherence as systems, and the empirical consistency with which certain characteristics tend to cluster together in certain classes of people.

Developers of typologies, however, rarely go beyond the process of identifying

and defining the various personality or patient types. The problem of determining whether a particular typology is meaningful or valid for some research or theoretical purpose is usually left unresolved. There are many reasons why the process of investigation stops short of pursuing the validity issue. On the one hand, there are persistent technical obstacles to translating the language of the typology into objective, quantitative terms so that further research can be carried out.

The second major source of difficulty in carrying the research forward is a conceptual one. How, for example, does one determine whether a typology is valid? What kind of criteria will serve to establish that the typology is meaningful from either the theoretical or the pragmatic point of view?

In our research, we have found it necessary to distinguish between two types of validity. The first can be called consensual, because it has to do with whether the patterns of behavior which are identified as types in one frame of reference can be verified by other observers in another situation. If the typology is validated in the consensual sense, then we are reassured concerning the reality of the system; i.e., the types are visible to other observers.

The second kind of validity to be considered and reported on in this paper is pragmatic in nature. The research will ask whether the system is predictive; for example, whether certain patient types will be demonstrated to be more responsive to treatment than other types. If this can be demonstrated, then it is possible to go on to the more refined clinical questions of whether certain types are differentially responsive to specific treatments.

It is with regard to problems in description and prediction that valid typologies would appear to have the most to contribute to theory and to the conduct of etiologic research. We have, in previous research, demonstrated that there is a diversity of behavioral and symptom patterns which characterize patients who are reliably diagnosed

¹ This paper summarizes the results of a series of studies. The author acknowledges the collaboration of Henri A. Lowery, Jonathan O. Cole, and Samuel B. Lyerly, in various phases of this research.

² Psychopharmacology Research Branch, National Institute of Mental Health.

as paranoid schizophrenia (Katz et al., 1964). The results of that study are a reflection of the current state of confusion which surrounds the concept of schizophrenia. Rumke (1960) has listed some ten currently unresolved contradictions concerning its etiology and manifestations. In the light of this confusion and the ambiguity of the concept of schizophrenia, it is difficult to see how any real progress can be accomplished in etiologic research without a meaningful and objective separation of the various kinds of patients who are currently assigned to that category.

With regard to predictive questions, it is clear, also, that to the extent that treatments are different in their effects, their success will be likely to vary as a function of the type of patient to whom they are administered. From the clinical standpoint, we are concerned with the specificity of a given treatment; whether it is likely to be effective in one class of patients as against another class. Again, such problems can only be meaningfully pursued in clinical research if methods for identifying and objectively separating out types of patients can be developed and applied.

Rationale and Procedure

Since exploring the validity of a system involves the selection of criteria, it is clear that in most research on the typology question the investigator has to come to terms with certain currently unresolved theoretical and measurement issues. Typological theory in the personality field is theory derived from Gestalt principles. Not only do we assume that the whole—the configuration—provides qualitatively different information than the sum of its parts, but, also, that a given trait draws its meaning from its context. The competing point of view holds that traits or symptoms are additive and that the whole can be reconstructed from the sum of its parts. This “holistic” way of viewing the organism appears to be congenial to clinical thinking, just as it had seemed to the holistic theorists

in the field of personality; e.g., Goldstein (1939) and Jung (1923), to be the most valid one for understanding the bases of behavior. It is the model upon which the clinician bases his assertion that the treatment response cannot be predicted from a single target symptom, unless one knows the context or configuration of characteristics in which the symptom appears.

It is clear that the additive theory is simpler in concept, and that from the standpoint of measurement it can be approached in a relatively precise manner. On the other hand, the “holistic” or “typological” point of view as a structural model creates untold complications for the measurement field. The answer to the question of the relative value of these approaches lies, however, in whether a particular typological approach leads to more information about the bases of behavior, is more conceptually satisfying, and is likely to be more powerful in predicting the subsequent response and behavior of patients than the simpler additive approach. In the research to be reported, the effectiveness of the typological model will be tested in relation to these questions.

The research steps which were followed in the development of and the attempt to validate the phenomenologic typology to be described in this report were the following:

(1) The selection of the basic content or the significant variables to be measured; i.e., the selection on theoretical grounds of that facet of the organism's functioning which is deemed to be highly significant from the standpoint of personality theory, and likely to lead to a classification of people (patients) which would have some generality across other areas of functioning.³

³ One notes, also, that the two points of view, the holistic and the additive, need not be at odds as regards the content of what it is that will be measured; i.e., they may agree on what aspects of the organism's functioning represent the significant variable with respect to personality. It is more a matter of how the information which is subsequently derived will be combined or integrated that is really the issue on which they differ. These differences in conception about personality structure lead to very different ways of dealing with the same basic data.

The primary content selected for measurement was the individual's pattern of social behavior.

(2) The development of a method for quantifying and separating out the various facets of the individual's functioning in this area (his social behavior).

(3) Determining whether a limited number of patterns (configurations) of behavior existed in a given heterogeneous population.

(4) Investigating whether the patterns, which are translated into types, can be consensually validated.

(5) Investigating the predictive validity of the typology.

The Content and the Development of the Method

The majority of the research which is aimed at describing and separating out patterns of symptomatology and behavior in the psychoses has dealt with their manifestations in the hospital setting. The effort to objectify and quantify observations of pathology as these occur in the hospital has not been extended to other settings, despite the fact that pathology is initially discovered and identified in the community. This situation has existed because it has been assumed that careful description and observation requires a controlled setting and the availability of trained observers.

The trained observers have generally been psychiatrists and psychologists who are viewed as the experts in the art of detecting the more subtle manifestations of pathology, and the semiprofessionals, the nurses, and ward attendants, who, because they work in the living situation of the patient, are in an advantaged position for the observation of his day-to-day behavior. The schedules of items of behavior which are provided these observers have grown out of various theories of psychopathology and out of experience with what it is in the patient's behavior that can be observed and rated reliably. The

outcome has been the development of sets of dimensions of pathology which are based on observations by the professionals of the patient's behavior in the interview situation (Wittenborn, 1951; Lorr et al., 1962) and other sets of dimensions (Burdock et al., 1960; Lorr and O'Connor, 1962; Goldberg et al., 1963), which are based on the analysis of observations in the ward situation. There should be some overlap of dimensions between these two situations, the interview and ward, but it would be very strange, given the difference in setting and types of observers, if there were complete agreement between the two. The general aim of the large majority of such research is the objective description of psychopathology in all of its manifestations.

The major gap in research in this area is the lack of objective information on the patient's behavior prior to entering the hospital. It is in the community that what is designated as psychopathology is initially manifested and identified. It is on the basis of such behavior that the patient is subsequently hospitalized, and despite the availability of detailed knowledge of his behavior following hospitalization, little knowledge exists of the relationship between behaviors in these two very different settings.

In the attempt to apply the descriptive-phenomenologic approach to the study of the psychopathology of schizophrenia, the authors felt that it would be salutary to depart from the expert's framework and to attempt, through the eyes of the lay observer, to reconstruct the nature of the behavioral pathology; to determine, for example, whether lay people would put the same kinds of behavior together that professionals do. Further, given the limitations of the expert observer; e.g., his theoretical preconceptions, restricted period of observation, and the limitations of the milieu in which the observations go on, it was felt that the key observer selected should ideally meet the following requirements: He

should (a) be in a position to observe the patient's behavior over a period of time in the patient's natural environment; (b) be naive about formal theories of psychopathology; (c) be somewhat detached emotionally as far as the patient is concerned so that his judgments could be viewed as reasonably objective.

It is unfeasible to locate any observer who meets all of these qualifications. In this research we selected the close relative of the patient and tried to construct the reporting scheme, a rating inventory, so that the influence of emotional involvement would be minimized.

The rating inventory, described in detail elsewhere (Katz and Lyster, 1963), is comprised of two kinds of items: Psychiatric symptoms which have been translated into lay language; and items which describe the nature of the individual's social behavior. The inventory is further characterized by the attempt to be as comprehensive as possible in its inclusion of items in the symptom and social behavior areas. It focuses as much as possible on concrete behaviors rather than on phenomena requiring judgments or inferences on the part of the observer. A format was used which would permit quantification and the development of a set of scores or a profile of the patient's behavior. It was intended that the inventory could be used both to describe a patient's behavior prior to entrance into treatment and, as part of a set of scales, to assess clinical and social adjustment at various intervals following treatment.

After a preliminary test found the a priori scales (e.g., level of "disturbed interpersonal behavior") to be generally valid in discriminating between extreme groups on the dimension of "adjustment," the inventory was administered to relatives of approximately 100 newly admitted psychotic patients at Spring Grove State Hospital in Maryland. The relative was asked to describe the patient's behavior just prior to his entrance into the hospital.

These data were used to determine, through a factor analytic method, whether a set of measures could be developed to provide a profile of symptomatic and social behavior. On the basis of this analysis, which was reported previously (Katz and Lyster, 1963), 77 of the original 128 items were assigned to 1 of 12 clusters which, from the cluster and a second order factor analysis, appeared to be measuring such factors as social obstreperousness, psychoticism (bizarreness and panic), withdrawal and depression, suspiciousness, and nervousness. The 12 clusters were then used to develop a formal scoring system for the inventory. These clusters are in table 1.

Development of the Typology

With regard to the development of a classification scheme, it was asked whether, within the fairly large heterogeneous population of patients who were classified as schizophrenics, relatives would be describing a limited number of patient types or behavioral patterns. Given the history of theory and experience with schizophrenia, there is, of course, every reason to believe that such behavioral subtypes exist. The questions were whether relatives, through their observational ratings, would separate out several distinct patterns, and whether the subtypes would be similar to those which had been previously identified in clinical research and practice. Because in the initial investigation we would have been hard pressed to predict the kinds of behavior patterns that relatives of psychiatric patients might perceive, we applied an empirical approach to answer the question as to what types (if any) existed in a sample of acute schizophrenics who were rated at the presumed height of the illness.

The empirical method adopted was to apply factor analysis (as an exploratory technique) to a matrix of the intercorrelations of the rating profiles of 24 acute schizophrenics who had been randomly selected from a larger sample of 100. This approach

TABLE 1.—KAS¹ Form R1: Relatives' Ratings of Patient Symptoms and Social Behavior—Subtest Clusters

(1) <i>Belligerence</i> ²		(8) <i>General Psychopathology</i>	
28.	Got angry and broke things.	5.	Acted as if he had no interest in things.
50.	Cursed at people.	12.	Felt that people did not care about him.
45.	Got into fights with people.	30.	Acted as if he had no control over his emotions.
113.	Threatened to tell people off.	31.	Laughed or cried at strange times.
(2) <i>Verbal Expansiveness</i>		32.	Has mood changes without reason.
100.	Shouted or yelled for no reason.	33.	Had temper tantrums.
106.	Talked too much.	34.	Got very excited for no reason.
99.	Spoke very loud.	42.	Bossy.
105.	Kept changing from one subject to another for no reason.	44.	Argued.
118.	Bragged about how good he was.	52.	Resentful.
(3) <i>Negativism</i>		55.	Got annoyed easily.
46.	Was not cooperative.	67.	Stayed away from people.
36.	Acted as if he did not care about other people's feelings.	71.	Preferred to be alone.
47.	Did the opposite of what he was asked.	73.	Behavior was childish.
48.	Stubborn.	79.	Very quick to react to something said or done.
56.	Critical of other people.	90.	Acted as if he were confused about things; in a daze.
51.	Deliberately upset routine.	91.	Acted as if he could not get certain thoughts out of his mind.
59.	Lied.	94.	Talked without making sense.
37.	Thought only of himself.	97.	Refused to speak at all for periods of time.
60.	Got into trouble with law.	98.	Spoke so low you could not hear him.
(4) <i>Helplessness</i>		110.	Talked about how angry he was at certain people.
93.	Acted as if he could not make decisions.	119.	Said the same things over and over again.
74.	Acted helpless.	121.	Talked about big plans he had for the future.
92.	Acted as if he could not concentrate on one thing.	127.	Gave advice without being asked.
3.	Cried easily.		
(5) <i>Suspiciousness</i>		(9) <i>Nervousness</i>	
40.	Thought people were talking about him.	20.	Got nervous easily.
107.	Said people were talking about him.	21.	Jittery.
43.	Acted as if he were suspicious of people.	38.	Showed his feelings.
108.	Said that people were trying to make him do or think things he did not want to.	22.	Worried or fretted.
(6) <i>Anxiety</i>		(10) <i>Confusion</i>	
19.	Afraid something terrible was going to happen.	85.	Lost track of day, month, or the year.
122.	Said that something terrible was going to happen.	86.	Forgot his address or other places he knows well.
18.	Had strange fears.	88.	Acted as if he did not know where he was.
111.	Talked about people or things he was afraid of.		
23.	Got suddenly frightened for no reason.	(11) <i>Bizarreness</i>	
125.	Talked about suicide.	116.	Talked about strange things that were going on inside his body.
(7) <i>Withdrawal and Retardation</i>		26.	Did strange things without reason.
76.	Moved about very slowly.	25.	Acted as if he saw people or things that weren't there.
8.	Just sat.	124.	Believed in strange things.
80.	Very slow to react.	24.	Had bad dreams.
70.	Quiet.	(12) <i>Hyperactivity</i>	
17.	Needed to do things very slowly to do them right.	7.	Had periods where he could not stop moving or doing something.
84.	Would stay in one position for long period of time.	13.	Did the same thing over and over again without reason.
		6.	Was restless.

¹ Katz Adjustment Scales.

² Items within a cluster are listed in order of importance for interpretation of the cluster. Order is based on part-whole correlations of individual items with the cluster.

is based on the method of locating clusters of people which was initially proposed by Stephenson as the Q procedure (1935). Our procedure differed from that described by

Stephenson in that the basic variables which were to be correlated were not separate personality items but standardized factors. These factors which are comprised of

groups of behavioral and symptomatic items, were described in the previous section.

Because a factor is more reliable and its meaning less ambiguous than a separate item, the correlations between profiles can be viewed as more stable and the bases for the relationships which exist between profiles can be more readily understood. The prime concern in this step of the analysis is that the basic units which are correlated are comparable (standard) with regard to their scalar characteristics, are both reliable and valid, and that information on these characteristics and the interrelationships among the variables, are known (Katz and Lyster, 1963). Since the Q procedure has been criticized on several counts (e.g., Zubin, 1964), I will indicate some of the thinking which went into the selection of the specific statistical techniques.

Although a number of approximate procedures exist for determining the similarity between profiles, for determining whether a limited number of homogeneous subgroups exist within a larger population, and for assigning subjects to categories, the selection of a particular technique represents a compromise based partly on statistical considerations, partly on the nature of the method of measurement being used, and partly on the content of the data which is being analyzed. For example, Cronbach and Gleser (1953), went into great detail on the mathematical bases and the comparative merits and limitations of the various quantitative similarity indices. The choice of correlation in this research as the measure of similarity between patients, was based on our judgment that given this type of method (an observational rating procedure) and this type of rating situation (each patient is rated by a different observer), the configuration, rather than the individual elements, represents the most relevant and reliable information with which to describe patients. The large number of profiles are derived from independent raters (relatives), each of whom

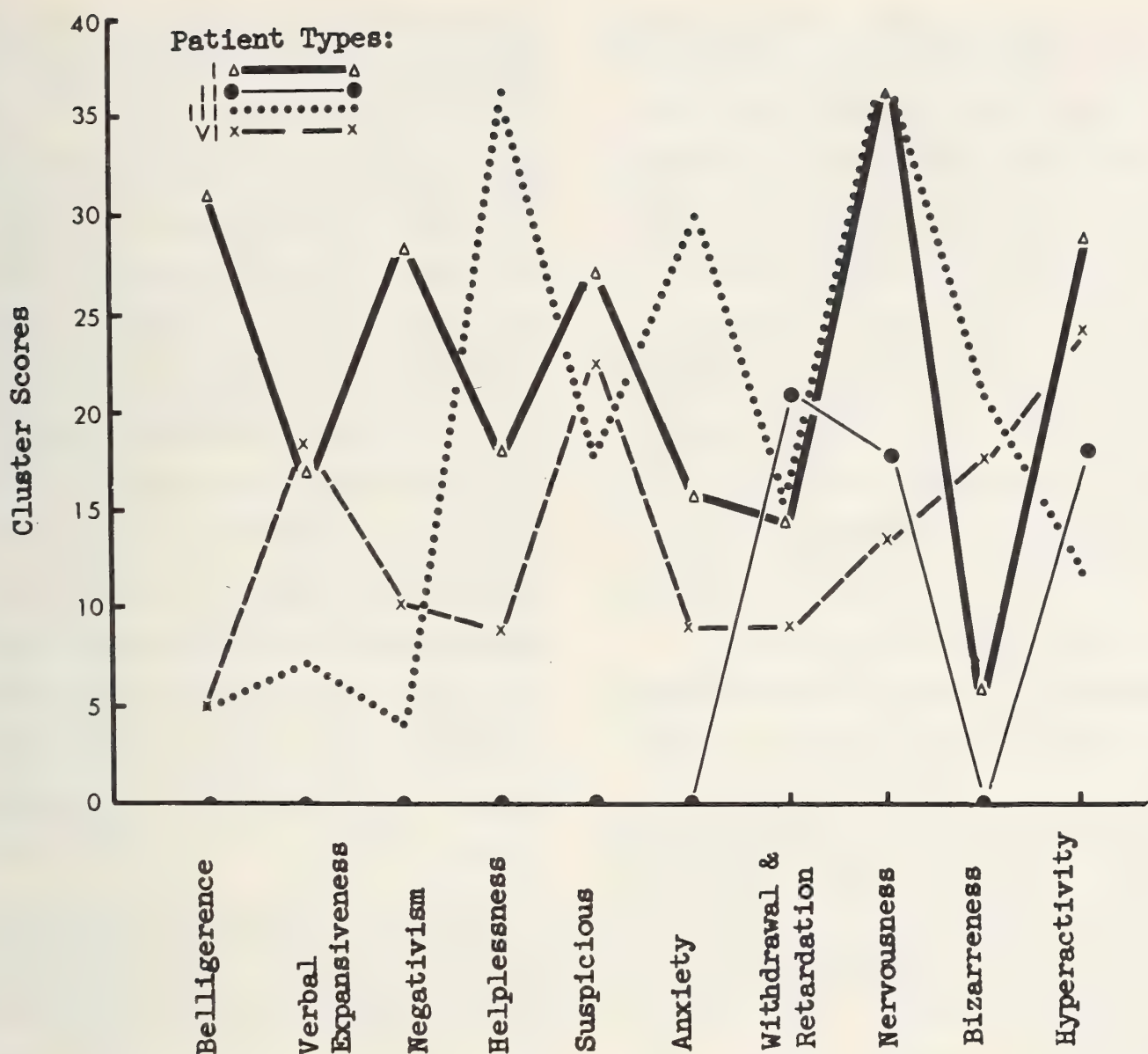
has his own standard or his own average level of ratings across variables. The data of interest is the rater's manner of comparing and contrasting behaviors in the patient, not the particular average level at which he rates. Similarity between profile shapes, then, best represents the basic information of interest here and correlation was judged to be the appropriate index.⁴

In the final analysis, however, the reality or utility of a classification system has to be sought outside of this kind of empirical separation of patterns. It is noted that the application of other statistical similarity indices and other clustering procedures might have resulted in slightly different types. Ways must, therefore, be sought to test a particular system's closeness to reality, its potentiality for making other phenomena understandable, and its utility.

The emphasis in this research on seeking ways to test the validity of the system grows partly out of the knowledge that exact mathematical solutions to these problems are not yet available and that our methods of measuring psychopathology are still at an early stage of development and only approximate in nature.

In the initial investigation the clusters of patients were found to be fairly discrete and whether the matrix was approached with the more precise principal components analysis or a simpler form of cluster analysis, the result was the same; i.e., sets of clusters derived from the two procedures were very similar. Six factors, "patient types," were found to account for the large majority of the variance in this matrix and these patterns were then delineated and found to be meaningful from a clinical standpoint; i.e., the patterns of behavior which are described by the six factors correspond to recognizable types of patients as

⁴ It would not be maintained, for example, that correlation is the best index to use in situations in which other types of methods, such as objective psychological performance tests, are used. A difference in level between intelligence test profiles has a different meaning and more significance than that same difference would have in the rating situation described above.



Clusters: Relatives' Ratings of Patient Symptoms and Social Behavior

FIGURE 1.—Four patient behavioral patterns derived from the typological analysis of relatives' ratings.

The Consensual Validity of the Typology

If the typology is valid, if the types identified in this analysis of relatives' ratings really are discrete, then other observers ought to describe the different types as having different patterns of characteristics. In attempting to investigate the consensual validity of the typology, we set the following questions:

(1) Will psychiatrists and psychologists, who subsequently see and rate these patients on the basis of a psychiatric interview, perceive the various types as manifesting different patterns of symptoms? If so, will the content and the pattern of the symptomatology described by the psychiatrists correspond to the description of the

patient type identified by the relatives? In other words, do the psychiatrists and relatives appear to be describing the same types of patients?

(2) Will the nurses perceive the different patient types as manifesting different patterns of behavior on the ward? If so, will the nurses and relatives agree with regard to the pattern of behavior presented by a given patient type?

In the "National Institute of Mental Health-Psychopharmacology Service Center Collaborative Study of Drugs and Schizophrenia" (1964), the availability of some 400 acute schizophrenic patients from 9 hospitals, who were rated by relatives at

the time of admission, made tests of these consensual validity questions feasible.⁶

The assignment of a patient to one of the types was accomplished by correlating (using the product-moment correlation) his profile with each of the six previously derived patient type patterns, and then assigning him to that category to which the profile-pattern correlation was highest. A patient was assigned to a type category only if the highest of the six profile-pattern correlations exceeded 0.40 and if it exceeded the next highest correlation by 0.10. In accord with these criteria, 61 percent of the total sample of patients were assigned to one or another of the six types.

All patients were interviewed during the first week of their hospital stay (as part of

⁶ The major requirements for the NIMH-PSC study were that the patients be newly admitted, acute, with evidence of symptomatology in two of seven major schizophrenic target symptom areas, and be between the ages of 16 and 41. The hospitals from which patients were drawn included State institutions, acute treatment services of city hospitals, and private treatment centers. The number of cases from each hospital ranged from 29 to 78, and the sample of 404 included 189 men and 215 women.

the NIMH study proper) and an Inpatient Multidimensional Psychiatric Scale (Lorr et al., 1962) was completed by two raters, usually a psychiatrist and a psychologist. The IMPS provides a profile of patient symptoms which include such factors as excitement, hostile belligerence, and paranoid projection.⁷ A profile of symptoms manifested during the initial clinical interview is thus available for each patient. Each patient's IMPS profile was assigned to that one of the six groups in which the patient was categorized in accord with the typological analysis of relatives' ratings. This results in six groups of profiles representing the six patient types. For each type in the relatives' typology, a mean profile of symptoms as seen independently by psychiatrists and psychologists can be derived. Four of these patterns (mean profiles) are presented in figure 2.

⁷ The admission interview IMPS data, which was available on all patients, was scored in accord with a system developed from a factor analysis of the item data on the 404 patients in the NIMH-PSC study. The rationale and the scoring system are described in Goldberg et al., (1963).

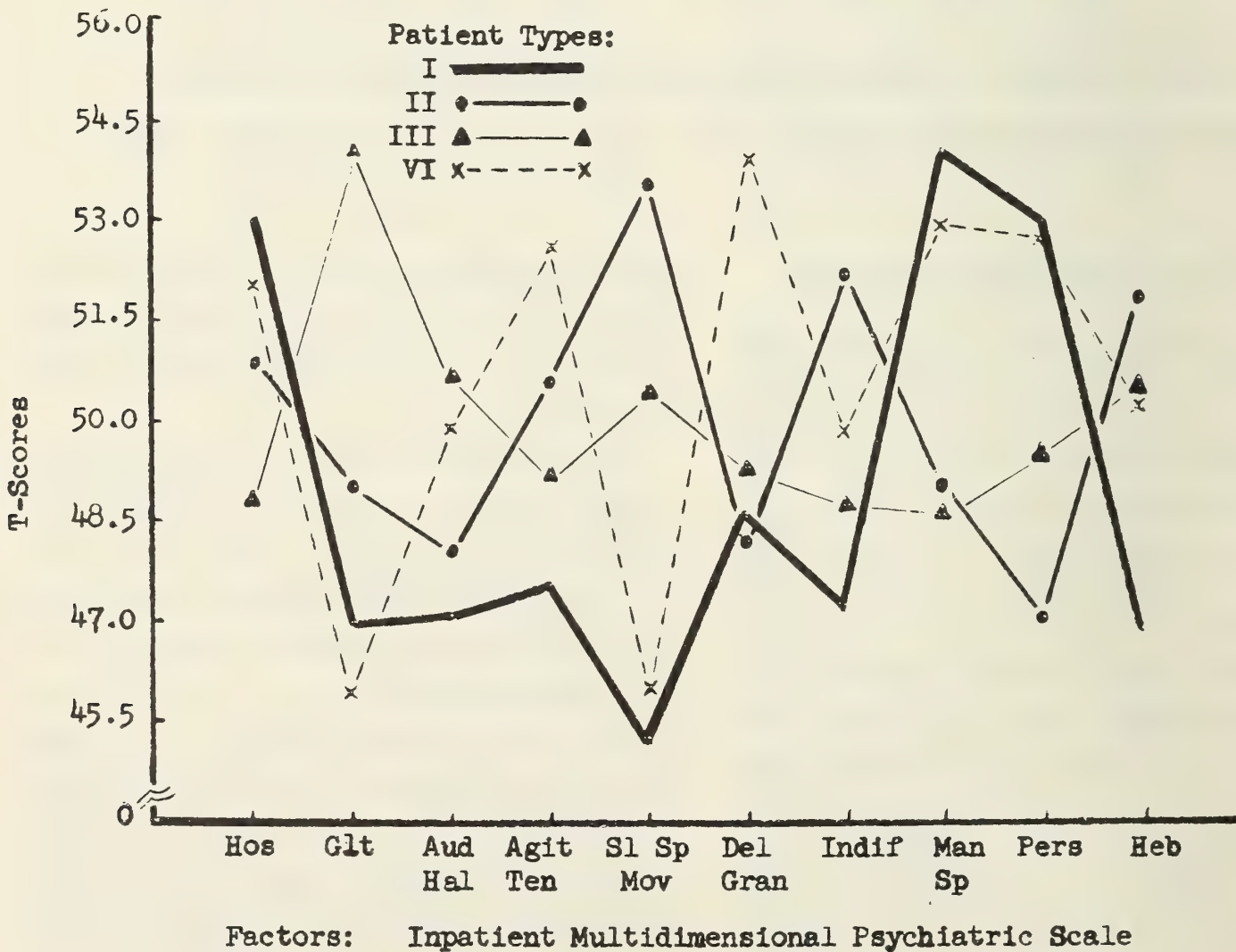


FIGURE 2.—Symptom patterns derived from psychiatric ratings of four patient types.

The first question asked was whether the several types would be seen as manifesting different patterns of symptoms by the psychiatrists and psychologists. The question of whether the six IMPS patterns are different in shape can be answered through an application of the multivariate analysis of variance. An approximation of this procedure developed by Greenhouse and Geisser (1959) was applied to the problem and the results are presented in table 3.

The interaction of patient types and measures will indicate whether the differences in profile shape are based on chance variation. The conservative test indicates that there is significant interaction beyond the 0.05 level of confidence. The profile shapes, in other words, are not parallel; i.e., the clinicians viewed the patterns of symptoms of the six types in the relatives' typology as different.

The answer to the question of whether the clinicians described the behavior of the types in ways similar to the relatives' description has to be arrived at on a qualitative basis. Since the items in the two rating scales, the relatives' ratings and the Inpatient Multidimensional Psychiatric Scale are different, the following conclusions are necessarily based on qualitative comparisons of the specific factors and the items in the two inventories. Compared on this basis, it can be seen that for four of the types the corresponding descriptions are

highly similar. The corresponding IMPS patterns are presented in table 4.

It is clear, for example, that the "well-integrated, hyperactive, belligerent, paranoid" group described by the relatives (type I), is seen by the clinician as prominent in "hostility, manic speech, and ideas of persecution," that the "withdrawn, periodically agitated type" (type II) is seen as predominantly "slower in speech and movement and apathetic," and the "agitated, expansive, bizarre paranoid" group (type VI) is prominent on "delusions of grandeur, manic speech, and ideas of persecution."

Type III turns out in this comparison to be differently described by the two sets of observers. Both described the depressed element ("helplessness" in the relatives' ratings, "guilt" in the IMPS), but the extreme anxiety seen by the relatives was not seen by the psychiatric raters. This may simply be a matter of perceiving these patients differently; it may be that the anxiety which existed prior to hospitalization is actually reduced following admission. It might also be partly a function of the lack of a factor in this IMPS scoring system which is comparable to the "anxiety, panic" measure on the relatives' scales. It should be noted, also, that the symptom pattern for type V only partially corresponds with the relatives' descriptions. Since the correspondence between the

TABLE 3.—Analysis of Variance of Psychiatric Symptom Ratings (IMPS Factors) of 6 Patient Types: Collaborative Study Sample

Source	d.f.	SS	MS	F
Scales (IMPS) ¹	9	8.55	0.95	0.01
Types.....	5	682.98	136.60	1.01
Individuals (within types).....	226	3,538.73	135.13	
Types X scales.....	45	11,865.29	263.67	² 2.86
Individuals X scales (within types).....	2,034	187,750.73	92.31	
Total.....	2,319	230,846.28		

¹ The majority of items in 3 IMPS factors, (disorientation, delirium, and memory deficit), were rarely, if ever, used in describing this sample of patients. This resulted in extremely low scale scores for the large majority of the sample and minimal variability. The 3 scales were, therefore, deleted from the analysis.

² Using the conservative test the F ratio (types X scales) is significant at the 0.05 level. Degrees of freedom required adjustment to 5 and 226 because of the lack of independence of the scale variables. See Greenhouse and Geisser (1959) for rationale for statistical procedure.

TABLE 4.—Psychiatric Symptom Patterns of 6 Patient Types Based on IMPS Mean Profiles: Collaborative Study Sample

PATIENT TYPE I (14 PERCENT) ¹ —AGITATED, BELLIGERENT, SUSPICIOUS ²		PATIENT TYPE IV (14 PERCENT)—WITHDRAWN, HELPLESS, SUSPICIOUS	
<i>High scores</i>	<i>Low scores</i>	<i>High scores</i>	<i>Low scores</i>
Manic speech. ³	Slower speech and move- ments. ³	Slower speech and move- ments. ³	Manic speech. ³
Hostility. ³	Hebephrenia	Auditory hallucinations. ³	Delusions of grandeur ³ .
Persecution. ³	Guilt	Indifference to environ- ment.	
	Auditory hallucinations		
	Indifference to environ- ment.		
PATIENT TYPE II (13 PERCENT)—WITHDRAWN, PERIODICALLY AGITATED		PATIENT TYPE V (10 PERCENT)—AGITATED, HELPLESS	
<i>High scores</i>	<i>Low scores</i>	<i>High scores</i>	<i>Low scores</i>
Slower speech and move- ments. ³	Persecution. ³	Indifference to environ- ment. ³	Hostility. ³
Indifference to environ- ment. ³	Auditory hallucinations.	Slower speech and move- ments. ³	Persecution. ³
Hebephrenia.	Delusions of grandeur.	Delusions of grandeur.	Guilt.
	Manic speech.	Hebephrenia.	
	Guilt.	Agitation and tension.	
PATIENT TYPE III (33 PERCENT)—ACUTE PANIC STATE		PATIENT TYPE VI (16 PERCENT)—AGITATED, EXPANSIVE, BIZARRE, SUSPICIOUS	
<i>High scores</i>	<i>Low scores</i>	<i>High scores</i>	<i>Low scores</i>
Guilt. ³	Manic speech.	Delusions of grandeur. ³	Guilt. ³
	Hostility.	Manic speech.	Slower speech and movement. ³
	Indifference to environ- ment.	Persecution.	
	Agitation and tension.	Agitation and tension.	

¹ Proportion of patient type in sample of 229 acute schizophrenics admitted to the 9 hospitals in the collaborative study.
² Titles refer to original types and not to IMPS patterns. See table 1 for more complete description of type.
³ One or more standard deviations above or below the profile mean for the patient types. Entries without a (³) are between 0.5 and 1 SD above or below the profile mean.

IMPS profiles and the six relatives' types derived from two separate studies is very good for all types except type V,⁸ this makes one suspect that this type is either less cohesive than the others, or responds in a more variable manner to hospitalization. As a further point, with regard to consistency between relatives' and psychiatric descriptions, it was noted that the IMPS data provided two factors, auditory hallucinations and ideas of persecution, which appear to measure two different manifesta-

tions of paranoid behavior. The item composition of the first factor describes paranoid symptoms which are based on fearful hallucinations; e.g., "hear voices that threatened punishment, torture, or death"; the items on the second, ideas of persecution, appear to describe a more hostile type of paranoid symptomatology with delusional ideas projected onto the real environment; e.g., "tends to blame, criticize, or hold other people, objects, circumstances responsible for his difficulties, failures, or frustrations." In the original typology there were three patient types which could be described as paranoid: The belligerent, manic (type I); the withdrawn (type IV); and the expansive, poorly integrated (type VI). It would have been expected that the psychiatric raters would have associated the hostile type of paranoid symptomatology (persecution) with the belligerent and the expansive paranoid

⁸ The other study referred to is one which was carried out earlier with the Research Department, Spring Grove State Hospital, through the cooperation of Dr. Albert Kurland and Dr. Mary Michaux. The study involved 200 psychiatric admissions to the hospital on whom relatives' ratings were available. The Spring Grove sample differed from the NIMH study sample in that it included both acute and chronic schizophrenics, and other nonschizophrenic but functionally psychotic patients. A more complete description of the characteristics of the sample is in Kurland et al. (1961). The results were very similar to those reported for the NIMH study. The types were found to be significantly different in their manifest symptom patterns. The NIMH study represented an attempt at cross-validating these earlier findings.

types, and the fearful symptoms (auditory hallucinations) with the withdrawn paranoid patient. As can be seen by comparing types in tables 2 and 4, this was borne out in the results.

With regard to question 2, a similar analysis was carried out using the set of measures derived from the nurses' ratings on the Ward Behavior Scale. The results of

the analysis of variance, which are presented in table 5, are similar to the results for the psychiatric ratings.

The interaction of types and measures is significant at the 0.01 level, indicating that the patterns presented by the types are not parallel; that they differ in shape. The patterns for the same four types are presented in figure 3. The prominent behavior

TABLE 5.—Analysis of Variance of Nurses' Ward Behavior Ratings (WBS Factors) of 6 Patient Types: Collaborative Study Sample

Source	d.f.	SS	MS	F
Scales.....	6	3.99	0.67	0.01
Types.....	5	1,817.79	363.56	1.82
Individuals (within types).....	222	44,278.99	199.46	
Types X scales.....	30	9,698.22	323.27	¹ 4.16
Individuals X scales (within types).....	1,332	103,623.05	77.80	
Total.....	1,595	159,422.04		

¹ Conservative test (5 and 222 degrees of freedom) significant at 0.01 level.

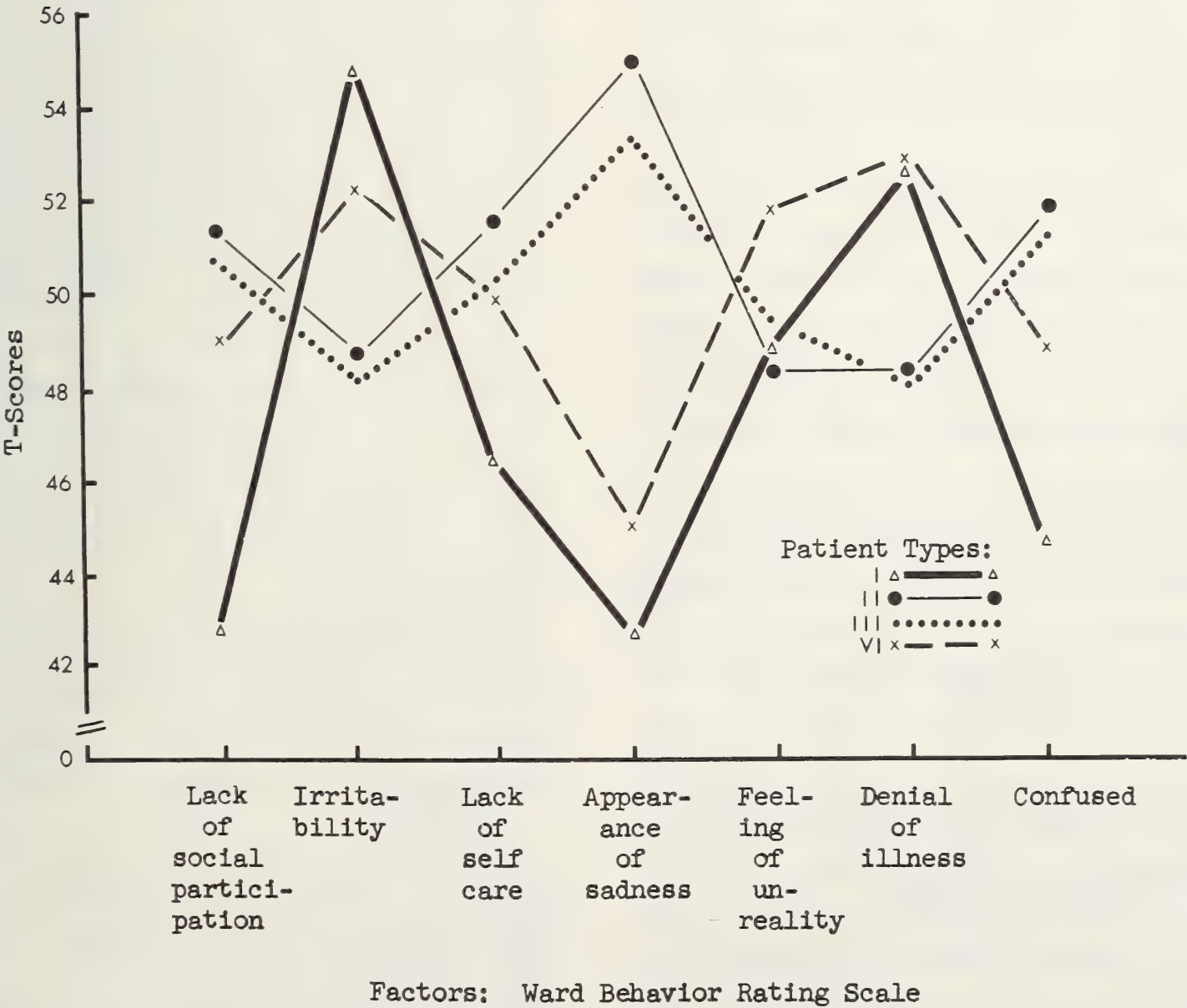


FIGURE 3.—Behavioral patterns derived from nurses' ratings of four patient types.

for type I, the “agitated, belligerent, suspicious” group is irritability and denial of illness. Type II, the “withdrawn, periodically agitated” patients are prominent on appearance of sadness and confusion. What distinguishes these comparisons from those of the psychiatrists and relatives in figure 2, however, is the fact that the pattern for type III looks very much like that for type II, and the pattern for type VI, except for a general heightening of symptomatology, specifically in the social participation and confusion areas, is similar to that for type I. The patterns still make sense from the standpoint of translating from the relatives’ to the ward nurses’ descriptions—but the discreteness of the types is not as distinct on the Ward Behavior Scales as on the IMPS. We would like to attribute this result in part to the fact that the factors which make up the Ward Behavior Scale are more global in nature, more experimental in quality, and not likely to be as sensitive to differences in patterns of behavior as the factors derived from the IMPS.

It might be expected, for example (after the fact), that since fewer factors are needed to account for the ward observers’ ratings, that the observer and the scale are likely to be less discriminating in separating out types of behavior. At any rate, the results indicate that the finer differences among the types are not discriminated by the ward personnel with this scale, although the more pronounced differences among the types—for example, the socially obstreperousness types (I and VI) versus the withdrawn or more depressed types (II and III)—are discriminated.

In summary, “patient types,” identified through an analysis of relatives’ ratings of the symptoms and social behavior of schizophrenic patients prior to their entrance into the hospital, turn out to manifest different patterns of behavior, one from another, when viewed by other groups of observers in other settings. These differ-

ent behavioral configurations are visible to a group of psychiatrists and psychologists, and to a significant, if less discriminating extent, to ward nurses. Despite their presumed differences in conceptual frameworks, despite the differences in measuring instruments used, and despite the variability introduced by using a large number of raters from nine different hospitals, the psychiatric raters’ descriptions of the behavior of the various patient types are highly similar to the relatives’ descriptions.

Predictive Validity of the Typology

Differential response of the patient types to drug therapy

It is clear from the consensual validity studies that the patterns of behavior manifested by the patient types are different, and that the descriptions of other kinds of observers agree quite well, in at least four of the types, with the patterns of behavior which were described initially by the relatives. This kind of validity is reassuring since it helps to establish the reality of the types, and argues against their being simply a function of the particular statistical procedures which were applied to the basic data.

The system, then, appears capable of objectively subdividing a large proportion of a population of functionally psychotic patients into relatively homogeneous subgroups who can then be described in terms which are meaningful to more than one type of observer.

The patterns of behavior which identify the various types based on relatives’ ratings are, for the most part, recognizable to clinicians as familiar types of patients. Although the terminology is somewhat oversimplified, the “belligerent, manic paranoid,” the “withdrawn, fearful paranoid,” the “withdrawn, periodically agitated” and the “acute, depressed, panic state” represent several variations of the schizophrenic psychosis which are meaningful to a good many clinicians. Nevertheless,

the categories are different from those previously described in the psychiatric literature, and we ask in the next study whether the typology will turn out to have theoretic and pragmatic value. A reasonable test of practical value of a classification system is whether it can predict the effects of psychiatric treatment.

It was possible to investigate this problem within the framework of the "NIMH-PSC Collaborative Study of Drugs and Schizophrenia" (1964). Patients in this study were randomly assigned on a double-blind basis to four treatment groups; three were treated with a phenothiazine drug (fluphenazine, thioridazine, or chlorpromazine) and one with a placebo, for a period of 6 weeks. Patients were initially assessed using several clinical rating instruments, which were based on observations of psychiatric interview and ward behaviors, and then reassessed for treatment effects after 1, 2, and 6 weeks of treatment. At the end of the 6-week period, the clinicians were also asked to rate the patients on a seven-point scale of improvement which extended from (1) "very much improved" through (4) "no change" to (7) "very much worse." This rating of global improvement was selected as the most appropriate clinical index and also turns out to be one of the criterion measures which discriminate between drugs and placebo.⁹

On the basis of previous clinical experience it might have been predicted that the tranquilizing drugs would be most effective with the belligerent, manic paranoid, the expansive, bizarre, poorly integrated schizophrenic, and the acute panic state, and would have been least effective, with the withdrawn, periodically agitated and the withdrawn paranoid types. The main con-

cern of this part of the study, however, was to determine empirically whether the six types would respond differently to drugs in general, or differentially to the several drugs. Since each patient was assigned randomly to a treatment (without knowledge of the type to which he belonged) this could be tested in the context of the NIMH study.¹⁰

As noted, 61 percent of the patients were assigned to one or another of six patient types. These were distributed among the six types as described in table 6. The criterion of treatment response selected was the clinician's rating of improvement ("global improvement scale") at the termination of the 6-week treatment period.

The data was subjected to a five (patient types) by three (drugs) analysis of variance which permitted a test of significance of the difference among types with regard to response to drugs in general ("patient types" main effect) and possible differential reactions of the types to the several drugs (interaction of "types" and "drugs"). Despite the large number of patients in the sample, it was necessary to exclude type V, because an insufficient number of patients were available in this category for the analysis.¹¹ Further, since the interest in this analysis was in differential response to drugs, the placebo group was not included. The results of the analysis are presented in table 6.

1. The significant main effect for "patient types" indicates that the groups respond differently to drugs in general.

2. The interaction of "patient types" and "drugs" was not significant, which indicates that there is no evidence, as far as

⁹ The result referred to is based on a discriminant function analysis in which the set of improvement indexes, psychiatric interview, and ward behavior ratings, were compared for their sensitivity in distinguishing between drugs as a group and placebo. The discriminant function was highly significant and one of the largest contributors to the separation of the two groups was the psychiatrist's rating of global improvement. (D. J. Clyde, personal communication, April 1964.)

¹⁰ It should be pointed out that patients were not assigned to types and then randomly assigned to treatment within each category.

¹¹ The sample was further reduced by the number of patients who had to be dropped from the study (following assignment) for administrative reasons; i.e., did not meet the criteria for inclusion in the sample, or because it was not possible to maintain them on drug treatment for the 6-week period. There was, also, a very small number (less than 3 percent) who improved so strikingly in the first week or two that it was necessary to release them from the hospital.

TABLE 6.—Comparison of the Different Patient Types on Response to Drug Treatment

Patient type	Mean global improvement ¹			Mean across 3 drugs	Rank order of response
	<i>Thioridazine</i>	<i>Fluphenazine</i>	<i>Chlorpromazine</i>		
I (N = 19), agitated, belligerent, suspicious.	2.12	2.14	2.75	2.34	4
II (N = 21), withdrawn, periodically agitated.	2.29	2.30	3.00	2.53	5
III (N = 48), acute panic state.	1.68	1.91	1.61	1.73	1
IV (N = 19), withdrawn, helpless, suspicious.	3.00	1.88	2.00	2.29	3
VI (N = 24), agitated, expansive, bizarre, suspicious.	1.83	2.00	2.08	1.97	2

¹ Low score means more improvement.

Source	Analysis of variance ¹		F
Between drugs.	d.f.	2	0.764
Between patient types.	4		² 3.099
Interaction (drugs X types).	8		1.644
Error.	116		
Total.	130		

¹ The method of unweighted means was used. (Winer, 1962).
² Significant at 0.05 level.

this general measure of improvement is concerned, that the “patient types” respond differently to the several drugs.¹²

The fact that the types do respond differently to the drugs as a group supports the contention of most clinicians in this field that the tranquilizing drugs are specific to certain symptom patterns rather than to schizophrenia, in general. The difference is significant at the 0.05 level of confidence; the five types are ranked in order of therapeutic response in table 6. The “acute, panic state” responds best; the “withdrawn, periodically agitated” group responds least to the drugs.

Multiple comparisons of pairs of “patient types” using the Newman-Keuls procedure (Winer, 1962) were carried out, and the difference between the latter two types is found to be significant at the 0.05 level.

¹² It should be pointed out with respect to this latter finding, that it has been very difficult to uncover any differences among the three tranquilizers, except in the area of side effects. There is an assumption that fluphenazine is more “stimulating” than the other two drugs, but so far little, if any, evidence from controlled studies has been found to support this opinion. It is expected that tests of the differences among the drugs which make use of a profile of specific measures of behavioral effects may demonstrate such differences or be capable of elaborating on the effects of these drugs in different types.

The differences between the other pairs did not reach a statistically significant level.

A question which could be raised is whether “therapeutic response” is a general characteristic on which the types differ and is not specific to drug treatments. One way of partially answering this was to compare the types on their reactions to placebo to determine whether they respond differentially to placebo and whether they assume the same rank order with regard to treatment response. A one-way analysis of variance was carried out using the improvement criterion, and the five groups were not found to respond differently to placebo. Further, the rank order of response to placebo was not the same as the rank order of response to drugs.

It has, also, been suggested that any relevant classification of patients might have accomplished the same results; i.e., would have demonstrated a differential response to drugs. We did in fact test out whether psychiatric diagnoses assigned during the first few weeks of hospitalization would predict response to drug treatment. Patients from the nine hospitals had been classified into the paranoid, catatonic, un-

differentiated, and schizoaffective categories and had been randomly assigned to the four treatments. It was possible, therefore, in the same kind of "patient groups by treatments" analysis of variance, using the global rating of improvement at 6 weeks as the criterion, to determine whether diagnosis was predictive of response. The results indicated that psychiatric diagnosis was neither related to response to drug in general, nor was there any evidence of any drug by diagnosis interaction.

It is clear, however, from these results, that the separation of the large population of schizophrenics into relatively homogeneous subgroups may permit a more refined analysis of the therapeutic effects of drugs. It turns out, as many have proposed in this field, that certain types of patients respond better to drugs than others. From a study of the results, it is apparent that acute panic states will respond better than withdrawn, periodically agitated patients. On the other hand, withdrawal itself need not predict a bad result, since the withdrawn, paranoid type responds better than might be expected. The belligerent paranoid does less well than the latter group; a result which also might not be predicted by many experienced clinicians.

We have had occasion, since this analysis, to cross-check these findings in a second NIMH collaborative phenothiazine study in which 8 hospitals and 448 patients were involved. In this second study, the patient population differed from the first in that it included a broader range of schizophrenia, both the acute and chronic disorders. Using the same criteria of assignment, 63 percent of 448 patients were assigned to one or another of six types based on the profiles derived from relatives' ratings which were carried out at admission. In this study, patients were randomly assigned to one of three phenothiazine treatments, acetophenazine, chlorpromazine, and fluphenazine, and rated on the global improvement scale at 5, 13, and 26 weeks.

The distribution of global improvement ratings at 5 weeks turns out to be very restricted in range (the large majority of patients improve and are rated at 1 to 3), which markedly reduces the sensitivity of the scale. Patient type III, the "acute panic state," and VI, the "disorganized, expansive, paranoid" type, again turn out to be most responsive, but the differences among types across the three drugs are not significant. Utilizing a more sensitive index of global change or improvement as the criterion, a measure based on a weighted composite of the changes in the IMPS and World Behavior Scale factors which occur between the initial and 5-week ratings,¹³ the differences in response among the six types turn out to be significant at the 0.05 level of confidence.

The order is the same as described above, the acute panic state and disorganized paranoid type being most responsive, with the differences among types being primarily a function of the marked superiority in response of type III when compared with the five other types. The least responsive in this analysis is I, the belligerent, well-integrated, paranoid type. The finding that a particular type of schizophrenic, the acute panic state, is most responsive to phenothiazines is one which has thus been confirmed in a second independent study.

The next step in this process was to ask a more refined question about the behavioral effects of the drugs in the different groups.

Patient Type and Treatment Specificity

Having established with the patient typology and a global measure of improvement that a certain behavioral type will do better on drugs than another type, several specific questions of clinical and theoretical relevance can be taken up through the applica-

¹³ The composite measure of change was developed by Nils Mattsson, Biometrics Laboratory, George Washington University.

tion of the typological approach. These have to do with whether—

- (1) given more discrete measures of change, one drug will turn out to be better than another for a given type of patient;
- (2) a particular symptom complex is affected differently by a drug as a function of the type of patient in which the symptom occurs.

In following through on these questions, it is necessary to turn to more discrete and more reliable indexes of improvement than the global rating. The global rating, despite its general value as an index of overall improvement, is obviously limited in terms of providing information as to the nature of the changes which take place in patients. These questions call for a comparison between patient types of change on a given symptom factor. Since it has been established however, that the patient types differ in the manner in which their illness is manifested prior to treatment, it is likely that the magnitude of initial differences, on any single factor, will obscure the meaning of such comparisons. We had, therefore, to seek compromise measures which would neither be too global, nor so specific that the major hypotheses could not be tested.

The measures selected for these studies were from Lorr's three second-order factors on the IMPS, which are comprised of combinations of the primary factors, and which he describes as the three basic facets of the general concept of morbidity. These are: Excitation-retardation; thinking and perceptual distortion; and schizophrenic disorganization. From the standpoint of measuring improvement, the selected measures assume that patients may get better in any or all of these three ways. Certain of the types can apparently be equivalent initially on a morbidity factor such as schizophrenic disorganization without necessarily being similar on certain of its specifics, such as the primary factor of motor disturbance.

There are, in the typology, two patient groups who are withdrawn and retarded and who are quite high initially in schizophrenic disorganization (which includes apathy and retardation) as rated by the psychiatrists. Type II is the "withdrawn, periodically agitated" and type IV is the "withdrawn, helpless, suspicious" group. In the NIMH study the three phenothiazine tranquilizers were presumed to differ slightly in their actions. One treatment, fluphenazine, was expected to be both stimulating and tranquilizing, while the other two were expected to be tranquilizing only. The more stimulating tranquilizer ought to be the better treatment for both of these types of patients. It turned out, as might be expected, to be difficult to test such questions when the research was not specifically designed for that purpose. For example, the small number of cases and a problem with the heterogeneity of the variances led to dropping the chlorpromazine group from the analysis. But it was possible to compare fluphenazine with thioridazine on the factor of schizophrenic disorganization in the two patient types. This factor was selected because the groups as noted were equivalent initially on it, and because it represents a major facet of the concept of severity of illness in schizophrenia.

The hypothesis to be tested is that fluphenazine will be more effective for the two patient types than thioridazine.

The results of the analysis of variance of posttreatment effects indicate that, contrary to expectations, the main effect of drugs is not significant. What is demonstrated is that the interaction of drugs and patient types is significant at the 0.01 level of confidence. (See table 7.) The conclusion would be that the effects of a drug differ for the two types of patients—or that the differential effect of two drugs on schizophrenic disorganization apparently depends on the type of patient to whom they are administered. The effects are illustrated in figure 4.

TABLE 7.—Comparative Effects of 2 Drugs on Schizophrenic Disorganization in 2 Withdrawn Patient Types

ANALYSIS OF VARIANCE¹: POSTTREATMENT RATINGS ON SCHIZOPHRENIC DISORGANIZATION

Source	SS	d.f.	MS	F
Types.....	145.65	1	145.65	1.01
Drugs.....	290.80	1	290.80	2.01
Types X drugs.....	1,158.68	1	1,158.68	² 8.02
Error.....	3,609.63	25	144.38
Total.....	5,204.76	28

ANALYSIS OF VARIANCE FOR SIMPLE EFFECTS

Source	SS	d.f.	MS	F
Drugs for type II.....	144.27	1	144.27	1.00
Drugs for type IV.....	1,305.21	1	1,305.21	² 9.04
Error.....	3,609.63	25	144.38
Types for Thioridazine.....	1,062.97	1	1,062.97	³ 7.36
Types for Fluphenazine.....	241.36	1	241.36	1.67
Error.....	3,609.63	25	144.38

¹ The method of unweighted means was used (Winer, 1962).
² Significant at the 0.01 level.
³ Significant at the 0.05 level.

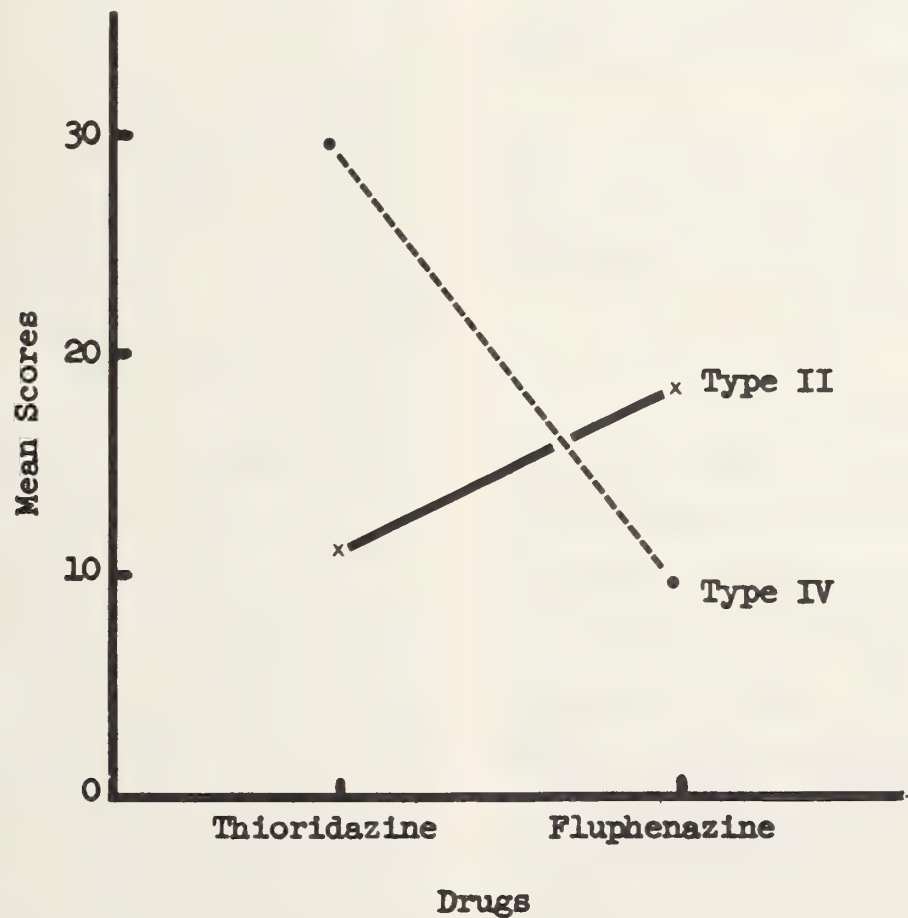


FIGURE 4.—Comparative effects of two drugs on schizophrenic disorganization in two withdrawn patient types.

Further analysis of this interaction indicates that:

(1) Fluphenazine was significantly more effective than thioridazine in reduc-

ing schizophrenic disorganization in the “withdrawn, suspicious” group (patient type IV).

(2) Thioridazine was significantly more

effective in the "withdrawn, periodically agitated" (type II) than in type IV.

Aside from demonstrating that a particular treatment is better for one type of patient than another, this analysis demonstrates how the same symptom complex can be differently effected by a treatment, as a function of the kind of patient in which the symptom occurs. The expected beneficial action of fluphenazine appears to be confined to one kind of withdrawn patient and not another. Why this is true is not clear, but the Gestalt and clinical principle that the behavioral attribute means different things as a function of its context appears to have support here.

In summing up—there is evidence:

- (1) That a given type of patient will respond better to tranquilizing drugs, in general, than another type.
- (2) That a particular drug will be more effective in one kind of patient than another drug.
- (3) That a particular symptom complex will be differentially effected as a function of the type of patient in which it occurs.

Concluding Comments

In reviewing the major issues which have to be resolved to carry out research of this nature, I noted that they are basically of two kinds, issues of theory and of measurement.

The rationale upon which the development of the typology was based was derived from the holistic theory of personality structure. The selection of social behavior as the primary content reflects a theoretical orientation which views such behavior as central to the concept of personality and as that facet of the organism's functioning which is most likely to lead to a meaningful classificatory system; i.e., a system which will have generality across other areas of functioning.

The experimental approach which was applied to the problems of identifying and separating out the various facets of social behavior was phenomenologic in nature. The approach differed from other approaches to the problem by using the quantified observations of lay people, relatives of schizophrenic patients, rather than psychiatric ratings, as the basic descriptive data for the development of the methods of measurement and the typology. The core issues to which the subsequent series of studies were directed were the consensual and predictive validities of the typology.

In this research it has been demonstrated that the patterns of behavior which characterize the several types can be distinguished by other observers in different settings. This was referred to as consensual validity because it represented evidence that the types which had been identified in one framework were visible to other kinds of observers in another setting. With regard to predictive validity, the research questions posed were whether the types would respond differently to drug treatment. It was found in one study and then confirmed in another that a type which was identified as the "acute panic state" does respond better to tranquilizing drugs than the other types. It could be further shown that if discrete, reliable measures of changes are used then two types who are similar in certain important respects, e.g., both "withdrawn," will turn out to respond differentially to different drug treatments. This occurred in a situation where the two drug treatments were not very different to begin with. This latter study should be repeated before any hard conclusions are drawn, but the study provided evidence for the holistic idea that the same "trait," or symptom complex could be differentially effected (by a treatment), as a function of the context in which the symptom complex occurs.

In addition to the results which bear on certain theoretical issues, the following findings and conclusions appear to be directly

related to current problems in clinical research:

1. There is strong evidence that there are several distinct varieties of paranoid schizophrenia. These paranoid types contrast markedly in their patterns of social behavior and symptomatology. These differences are most pronounced in the manifestation of belligerence and withdrawal but are, also, reflected in the content and the actual quality of paranoid thinking. The dissimilarities among these types, when they are viewed either from the characterological or the behavioral standpoint are highly visible, and are found to have implications for psychiatric treatment. The evidence for the existence of these types is not only from the recent work on typologies, but, also, from the research on current deficiencies in the diagnostic system referred to earlier (Katz et al., 1964) in which it is clear that highly diverse, but discrete kinds of patients are being included in the same diagnostic category.

2. In future studies we will seek out ways to clarify and refine the definition of the acute panic state since this type of patient turns out to be the most responsive to tranquilizer treatment. From the findings in the consensual validity studies it is possible to characterize their behavior following hospital admission, but the picture, especially with regard to possible changes in the level of anxiety as a function of hospitalization, is still not very clear.

The fact that types are differentially responsive to drugs merely confirms the notion that the field will probably progress more rapidly in clinical and etiologic research, if the initial stages of research in psychopathology, i.e., description and classification, are more carefully attended to. A rationale for a phenomenologic approach to the problem of classification has been presented. It is not suggested that objective and refined measurement of overt psychopathology and social behavior will, in themselves, solve the etiology problem. It is expected, however, that the likelihood of

locating uniformities in other important areas of functioning will be greatly enhanced—and thus, the prospects for progress in etiology increased—when a reliable descriptive and classificatory framework is available in the field of psychopathology.

REFERENCES

- BURDOCK, E. I., HAKEREM, G., HARDESTY, A. S., and ZUBIN, J. "A Ward Behavior Rating Scale for Use With Mental Hospital Patients." *J. Clin. Psychol.*, 1960, 16, 246-247.
- CRONBACH, L. J., and GLESER, GOLDINE C. "Assessing Similarity Between Profiles." *Psychol. Bull.*, 1953, 50, 456-473.
- GOLDBERG, S. C., COLE, J. O., and CLYDE, D. J. "Factor Analysis of Ratings of Schizophrenic Behavior." *Psychopharmacology Service Center Bulletin*, 1963, 2, 23-28.
- GOLDSTEIN, K. "The Organism." New York: American Book, 1939.
- GREENHOUSE, S. W., and GEISSER, S. "On Methods in the Analysis of Profile Data." *Psychometrika*, 1959, 24, 95-112.
- JUNG, C. G. "Psychological Types: Or the Psychology of Individuation. (Translated by H. Godwin Baynes) New York: Harcourt Brace, 1923.
- KATZ, M. M., and COLE, J. O. "A Phenomenological Approach to the Classification of Schizophrenic Disorders." *Dis. Nerv. System*, 1963, 24, 147-154.
- KATZ, M. M., and LYERLY, S. B. "Methods of Measuring Adjustment and Social Behavior in the Community: I. Rationale, Description, Discriminative Validity, and Scale Development." *Psychol. Reports*, 1963, 13, 503-535.
- KATZ, M. M., COLE, J. O., and LOWERY, HENRI A. "The Nonspecificity of the Diagnosis of Paranoid Schizophrenia." *Arch. Gen. Psychiat.*, 1964, 11, 197-202.
- LORR, M., and O'CONNOR, J. P. "Psychotic Symptom Patterns in a Behavior Inventory." *Educ. Psychol. Meas.*, 1962, 22, 139-146.
- LORR, M., MCNAIR, D. M., KLETT, C. J., and LASKY, J. J. "Evidence of 10 Psychotic Syndromes." *J. Consult. Psychol.*, 1962, 26, 185-189.
- "NIMH-PSC Collaborative Study Group: Effectiveness of Phenothiazine Treatment of Acute Schizophrenic Psychoses." *Arch. Gen. Psychiat.*, 1964, 10, 246-261.
- RUMKE, H. C. "Contradictions in the Concepts of Schizophrenia." *Compr. Psychiat.*, 1960, 1, 331-337.
- STEPHENSON, W. "Correlating Persons Instead of Tests." *Charac. and Pers.*, 1935, 4, 17-24.
- WINER, B. J. "Statistical Principles in Experimental Design." New York: McGraw-Hill, 1962.

WITTENBORN, J. R. "Symptom Patterns in a Group of Mental Hospital Patients." *J. Consult. Psychol.*, 1951, 15, 290-302.

ZUBIN, J. "Taxonomy in the Mental Disorders: A Historical Perspective." Paper presented in symposium: "Explorations in Typology With Special Reference to Psychotics." At the American Psychological Association, September 1964.

Supplementary Comments by Dr. Katz As Presented at the Conference

In this brief presentation, rather than summarize the contents of the formal paper, I will explore more fully the rationale, the theory upon which this research is based. I do this partly in response to some of the critical discussion of the paper, and partly because I think that in emphasizing methodology in the formal paper—that I have neglected to present some of the ideas which underlie the work as a whole.

I will indicate the assumptions which guided the research and then discuss these assumptions more fully.

(1) The theoretical model upon which this work is based is the typological; it holds that personality differences are not merely quantitative, and that people can be classified into qualitatively different types.

(2) The most effective way of distinguishing among these types is through the observation and measurement of their patterns of social behavior.

(3) The differences in behavioral patterns between the types are exaggerated and thus more easily identifiable in the mental disorders.

(4) The methods of measurement and the statistical approaches which are applied to the analysis of the patterns of behavior should be appropriate to the theoretical model of personality structure—in this case, a typological model.

(5) The most important test of the validity of the typology is one which determines whether the derived types actually exist, whether they can be identified in some

framework other than where they were initially uncovered.

(6) To the extent that the typology is valid and that the portrayal of the types are accurate, the types are likely to react differently to treatments which markedly affect behavior.

Determining whether a typology will predict response to treatment is, then, a secondary goal of the research. Effective prediction depends on whether the basic theory, the typology, is valid, and whether it has been represented accurately in methodological terms.

The concept of qualitative types has a long and controversial history in personality theory, and this history was explored and extended by Carl Jung in his classic treatment of the subject, "Psychological Types" (1923). Because this work has so strongly influenced the direction and quality of research in psychopathology and contributes so importantly to our understanding of results in this area, I will briefly relate his thinking on the matter.

Jung believed that there were two major types of personalities, the object oriented and the subject oriented. In the one case, the orientation involved an outward interest toward the object, and in the other, a movement of interest away from the object, toward the subject and his own psychological processes. The two types might be distinguished as a function of the strength with which each is pulled toward the outer object in his world. It is not so simple in actual practice to identify these types because as a matter of theory, Jung later found it necessary to differentiate four subtypes within each major type. This subtyping was based on the predominance with which a particular psychological function is relied on; e.g., thinking, perception, etc. Further, with regard to the problem of identifying the basic types, there are complications introduced by the culture; the manner in which these attitudes will actually be manifested is obviously shaped by the characteristics of the particular culture.

If all of these complications were not sufficient to discourage those who might attempt to identify and to separate the basic types—he also pointed out that what he was describing were typical attitudes—and as such the predominance of one of the basic mechanisms is merely relative—that, as a matter of fact, both exist in the same person. In “Psychological Types,” Jung was able to demonstrate that these two basic types have been identified throughout human history by philosophers, poets, and psychologists. He marshals a great deal of evidence that all of the apparently diverse formulations which arose from these varied orientations had a common basis, and it is only because of this common basis that we find not only that the idea that there are qualitatively different types of personalities is still being pursued, but that contemporary typologies based on current theory are not very different than those delineated in his classic treatment. In current research the use of classifications like field dependent versus field independent—outer directed versus inner directed—manic versus schizoid, provide cause to wonder whether the basic ideas in this area have really changed very much at all.

These ideas of Jung and those who preceded him underlie a great deal of the way we still think and develop theory in psychopathology and personality.

I noted in my research report that our series of studies involved developing a multivariate method for measuring social behavior, and then fitting certain statistical approaches to the data in order to separate out subtypes within schizophrenia. It was noted that the approach to locating types was empirical in nature. The types were identified in an unbiased manner within a larger heterogeneous population and then each is described on the basis of the particular configuration of social behavior which is manifested. Now if these empirically derived types are valid, if they are very likely to exist, then it will be useful to look at and attempt to interpret the types

in terms of Jung’s theory. If they make sense in that framework, then something is gained in terms of further thinking in this area by creating this bridge between a particular theory of personality and the manifestations of psychopathology.

If they correspond fairly well with what has been described, it is further support (of a theoretical nature) for the validity of the typology. Before attempting, however, to integrate the characteristics of our system with existent theory, it is necessary to look outside of the original framework within which the types were identified for confirmatory evidence that they actually exist. It was for this reason that we turned to observers in other settings to determine whether the patterns of behavior which characterize the types are visible to others. The evidence for consensual validity from several sources was presented in the formal report. I have on the basis of that evidence relooked at the patterns to determine whether they appear to be reflecting the basic personalities described by Jung. I think several of them reflect the basic personality orientations described, but because there obviously is no direct test of that proposition, I consider the comparisons exploratory. It could be that other ways of interpreting the differences between the types would be more productive.

I would like to refer you to one graph, however, from the consensual validity work which demonstrates quite vividly the contrast between the patterns of behavior manifested by certain of the patient types. (See fig. 3.)

These four patterns reflect the kinds of ward behavior which are prominent in four of the types. Type I (the heavy line) is the belligerent, hyperactive (Jung’s outgoing type), and type II (the light line with the circles) is the withdrawn, more catatonicleike schizophrenic. Their patterns of behavior with regard to social participation, irritability, sadness or apathy, and denial are consistently opposite and exaggerated and could be interpreted as characterizing

the two basic personalities, the outgoing and the inwardly turned psychoses.

Just a final point with regard to this problem of predictive validity: I noted that if the typology is valid, that if it can be tied to basic psychopathology or personality theory—then it is likely to be more effective in the long run in predicting the effects of any definable treatment—or predicting other behavior—or the effects of other situational influences on the patient. To the extent that these types are really different, they ought to respond differently to treatments which have demonstrated effects on manifest behavior. Evidence that this actually happens is presented in the results of my research report—but I thought it important to indicate why this successful prediction result is dependent on or secondary to the development of a valid typology—and why the methodologic and statistical approach is tied to the theoretic notion of types—and not to the problem of predicting the effects of a particular kind of treatment.

Discussion

GEORGE KELLY, Ph. D., *Professor of Psychology, Brandeis University, Chairman*; ARDIE LUBIN, Ph. D., *Research Psychologist, U.S. Navy Medical Neuropsychiatric Research Unit, San Diego, Co-Chairman*

(Papers: "The Phenomena of Depressions," Roy R. Grinker, Sr., M.D., Department of Psychiatry, Michael Reese Hospital, [P. & P.I.], Chicago, and Jum C. Nunnally, Ph. D., Professor of Psychology, Vanderbilt University, Nashville. "A Typology for Functional Psychotics," Maurice Lorr, Ph. D., Chief, Outpatient Psychiatric Research Laboratory, Veterans Administration, Washington, D.C. "Studies of Quantitative Approaches to Psychiatric Classification," John E. Overall, Ph. D., Director, Research Computer Center, University of Texas Medical Branch, Galveston, and Leo E. Hollister, M.D., Veterans Administration Hospital, Palo Alto. "A Phenomenological Typology of Schizophrenia," Martin M. Katz, Ph. D., Chief, Special Studies Unit,

Psychopharmacology Service Center, National Institute of Mental Health.)

FORMAL DISCUSSION OF PAPERS

	Page
ARDIE LUBIN, Ph. D.	322
SAMUEL GREENHOUSE, Ph. D.	328
HEINZ E. LEHMANN, M.D.	330

OPEN DISCUSSION

Drs. Zubin, Grinker, Nunnally, Lorr, Overall, Katz, Ball, Hamilton, Klerman	344
--	-----

Discussant's Remarks

ARDIE LUBIN, Ph. D., *Research Psychologist, U.S. Navy Medical Neuropsychiatric Research Unit, San Diego Naval Hospital, Calif.*

The four papers presented here, on the methodology of developing new typologies, cluster together very well in terms of their presenting symptoms, dynamic etiology, and temporal transactional sequence. In each case, the first statistical procedure was to carry out a factor analysis. In fact, several factor analyses usually were done. These factor analyses generally provided linear combinations of the original scores or ratings. The linear combinations were then used to compute correlations and covariance matrices between selected patients (although in one case, the factor analysis was abandoned and the original scores were used).

Next the intercorrelations between the patients were attacked with variety of weapons; cluster analysis, powered vectors, higher order correlations, congruency coefficients, etc., more than one of these methods generally being applied. In this way, clumps of patients were hacked out of the correlation matrix, and a new typology was born.

This period of colossal computation was followed in most cases by a certain amount of skepticism as to the usefulness of the

results. One paper projected a definite depressive mood; the following are quotes from the Grinker and Nunnally paper.

Feelings of hopelessness:

• • • the major problem is that no one can say what class of decisions is most relevant to a scheme of diagnosis.

Feelings that the world should be changed:

• • • unfortunately people are not sympathetic to the mathematical requirements for employing such types because most people prove to be atypical.

And finally a feeling of helplessness:

In order to perform the necessary experiments, however, it would be required either to assign each patient to all treatments or to assign randomly each patient to the one treatment, neither of which is practicable!

These quotations from the Grinker and Nunnally paper tend to be rather pessimistic.

Lorr has some feeling of incompleteness about the typology he has created but is much more optimistic, in terms of the future.

• • • the types will be of little interest unless and until some evidence can be presented of their validity • • • . Do the types respond differentially to available psychiatric treatment modalities? Are they useful for predicting length of hospitalization, duration of illness, or treatment response? • • •

However, Lorr felt no difficulty about testing the question of prognostic and therapeutic validity. In his table 9 some differential therapeutic outcomes are listed for his types. Unfortunately he had no time left for statistical assessment, so we do not know if the Lorr types have therapeutic validity.

From the very beginning of his paper, it is clear that Katz has the validity question firmly in hand. In particular he proposes that new typologies be tested for consensual validity and predictive validity. Consensual

validity refers to the question of whether psychiatrists, psychologists, nurses, relatives, etc., agree with one another. Predictive validity is what I have earlier referred to as therapeutic validity. (I feel that the whole question of agreement and disagreement on patient classification is irrelevant, immaterial, and highly prejudicial to getting on with the job. If a typology is useful we can always find objective reliable ways of sorting the patients into the appropriate categories.)

Katz points out that the crucial test is whether the interaction of patient-type-by-drug is significant. Do the drugs have the same rank order of effectiveness in every group of patients? Unfortunately, Katz finds the answer "Yes," and that makes this typology therapeutically invalid as far as differential treatment is concerned.

However, Katz reasoned that for at least two of his types, "schizophrenic disorganization" was a better measure of improvement than global score. The chlorpromazine treatment group was dropped from the analysis because of the small number of cases and heterogeneity of variance. (I think this is a wholly unnecessary ritual sacrifice to a mythological statistical rigor. Here and later in the use of the unweighted means anova, a great many subjects were wasted. When the cost of each subject is considered in terms of money, energy, and time, it is not just bad statistics to discard available information, it is a sin!)

The type-by-drug interaction was significant. An examination of Katz's figure 4 shows that this interaction is most useful—it is of the disordinal type, the rank order of the drugs differs for each type. So at last Katz has shown that even in the restricted field of phenothiazines, his typology has therapeutic validity; it enables the physician to use different treatments for different kinds of patients.

Overall and Hollister also had no difficulty in locating studies where patients had been assigned randomly to each treatment.

Never underestimate the power of NIMH! This allowed O and H to test their typology for therapeutic validity.

First they used a distance function (multivariate) to allocate patients into the standard diagnoses; paranoid, depressive, and schizophrenic. Then the type-by-drug interaction was used, as in Katz's study, as a criterion of validity. Seven drug treatments were used and a significant disordinal interaction was found. The most effective drug for paranoids was triperidol. Oxyperline gave the best results for both schizophrenics and depressives. This shows that the usual psychiatric distinctions can and do have therapeutic validity. To me, this now throws the burden of proof on any new typology. Can it be shown that the new typology incorporates all the therapeutic validity of the standard psychiatric typology and adds new validity? This question must be answered affirmatively before the standard classification scheme is discarded for any new nosology.

The next O and H study was just such a test. Carefully classified depressive patients were subjected to empirical profile analysis and three types emerged: Anxious, hostile, and retarded. All of these types are homogeneous from the standpoint of standard psychiatric classification. The patients were assigned randomly to a thioridazine and a imipramine treatment group. A significant type-by-drug interaction was found, again of the useful disordinal type. Thioridazine was best for the anxious group, imipramine was best for the retarded group. The hostile patients improved no matter which drug they took. Thus they can conclude that their empirically derived types added new therapeutic validity to the standard psychiatric classification of depressed.

Perhaps the most important study performed by Overall and Hollister was their creation of a typology based on drug response. Since there were seven drug treatment groups, they were able to start with seven types; i.e., the best responders in each treatment group. The use of pretreatment

ratings in a discriminating function analysis collapsed the seven types into four types. Methodologically speaking, this is an important innovation since classification by drug response maximizes the chance of finding a typology that will have therapeutic validity. No comparison of this typology with standard psychiatric diagnoses was attempted so it is not clear whether the drug-derived nosology made a uniquely valid contribution to the treatment-by-type interaction.

A Diagnostic Dialog With a Drug-Devised-Typology-Devotee¹

The DRUG - DEVISED - TYPOLOGY - DEVOTEE (DDTD). Overall and Hollister specifically say that a completely adequate system for psychiatric classification can not be based solely on patient responses to current therapeutic techniques. Why not? Isn't the sole purpose of diagnosis to help treat the patient?

The CRITIC (C). Therapeutic significance is highly valuable but it is only one of the functions of classification. In general medicine, classification takes into account the etiology, the pathological process, the mode of spread, the prognosis, the pattern of symptoms, and the response to treatment.

The DDTD. Once you have decided to derive a typology from objective data, is there a better alternative to the treatment-response (T-R) analysis? Even if you do find out later on that the etiologies of the patient maladies are mixed, that the outcome varies, the mode of spread does not appear to be constant, etc., is there any doubt about the usefulness of the T-R types for the physician?

The C. A question for a question. Would you consider it important to classify together headaches, rheumatism, and bruises because they all respond to aspirin? In the

¹ I am indebted to Dr. Max Hamilton of the University of Leeds for some of the caustically critical comments given here.

same way, what about the fact that both meningococcal meningitis and syphilis respond to penicillin? Basically we will probably progress faster and further if we stick to the nosological procedures that have been so effective in physical medicine. Remember that there are basic questions of biochemistry, genetics, neurophysiology, etc., involved in psychiatric classification. If we insist on a purely pragmatic approach, psychiatry will never be able to root itself firmly in the bases of biological science.

Now that I have shown you the futility of using only the T-R analysis, tell me why you scorn profile analysis, factor analysis, and other statistically sound methods of deriving types.

The DDTD. Sound? In what way? As long as we have no way of deciding which classification systems are optimal, there is no limit to the number of such methods that can be created. Of the devising of cluster analysis methods there is no end. Can any of the statistical methods of typologizing used at this conference, when applied to the data of Linnaeus, result in the phylogeny of Linnaeus? Can any of them duplicate what we now regard as the "true diagnoses" of physical medicine? No? Then I say consign them to the flames for they can contain nothing but sophistry, illusion, and huge computer bills.

The C. We can disregard your last remark since you yourself have cited the therapeutic validity of the empirically derived types of Overall, Hollister, Katz, and Lorr. Furthermore it is not true that there is no way but treatment-response to decide on the optimality of a classification system. In almost all cases, profile clustering methods do have some quantitative criterion that they are trying to maximize. In principle, this will always lead us to the optimal typology based on empirical data.

The DDTD. Yes; but each kind of cluster analysis uses a different criterion as the one

to maximize. So each kind of analysis may converge on a particular typology, but they will not approach the same unique diagnoses, no matter how diligently pursued. And what makes statistical processing of empirical data so important? Loewi and Kekulé did pretty well just using their dreams.

What we need in psychiatry are some good dreamers!

The C. You really can't say that the empirical methods of typologizing won't converge because we don't know yet. And remember, your T-R typology must be based on empirical data also. How do you propose to treat your data so as to avoid the difficulties you criticize in factor analysis, cluster analysis, etc.? Just what criterion would you maximize?

The DDTD. My basic criterion is that patients be assigned to treatments in such a way that the sum of improvement scores be a maximum: that there be no other way of assigning patients that will produce a better overall improvement score. When patients are assigned to treatments on the basis of their classification, this implies that the treatment-by-type interaction should be maximized. Imagine that we have N patients and k treatments. Further suppose that each of the patients has been subjected to a treatment, the effects of the treatment wiped out so the patient is returned to his original level, then another treatment administered, wiped out, etc., until the patient has experienced every treatment. The result would be a N by k matrix of improvement scores. For each subject, find his highest improvement score and assign him to the treatment that produced that maximum improvement. In this way we create a maximum of k types. Regrouping the patients by types, we now can calculate a treatment-by-type interaction which presumably will be the maximum obtainable. (Statisticians will recognize this as the simplex solution to the allocation-of-trans-

portation problem, when quotas do not exist.)

But, as Grinker and Nunnally have pointed out, no such multiple-treatment design could or should be carried out, even with the aid of divine intervention, or the federal government. Of course the Overall-Hollister method could be used, and improvement scores within each treatment group used to generate the types. However, it may be possible to make use of more information than just the drug response. Suppose for each treatment group we predict the response to treatment from pretreatment data. Relative success in this has been reported by Jonathan Cole and others in more distant portions of the forest. So we can predict the response of every patient to every treatment and in this way fill in the N by k matrix of improvement scores.²

The C. Yes; if you will look at page 7³ of the O-H paper, the use of multiple regression is advocated for predicting treatment response. But, as one who has used multiple regression, I can assure you that it is most unsatisfactory. Linearity is demanded; intricate relations such as Saunder's moderator variables and the interactions of analysis of variance are entirely disregarded. Your predicted improvement scores will be stripped of the real complexity contained in your pretreatment data.

The DDTD. Recently it has become clear to some of us, such as John Plag, my colleague at the U.S. Navy NP Research Unit, San Diego and R. J. Wherry, Jr., of the U.S. Navy Aviation School of Medicine at Pensacola, that analysis of variance really is based on a linear model: i.e., that multiple regression appropriately handled can take care of qualitative predictors, non-linearity, moderator variables, interactions—you name it. Statisticians seem to have

known this all along. For example, see M. G. Kendall's least-square solution to analysis of variance with disproportionate frequencies. It is straight multiple regression plus the coding of all efforts into what R. J. Wherry, Jr., calls pseudodichotomous variables and most statisticians call binary or dummy variables. Even I knew the answer back as far as 1951 (A. Lubin, and J. May: "Eliminating Superfluous Variables in an Analysis of Variance With Disproportionate Frequencies." *Biometrics*, 7,118). It is only now that I know that I didn't know what I knew then.

The C. Please stop confusing me with lamentations about what it is that we know now that we never knew we knew before. I will let you assume for the moment that multiple regression can handle the complex pretreatment patterns that I think are relevant. But how do you know that you have the right improvement score? Remember that Katz had to shift his definition of improvement to get a significant type-by-treatment interaction. And that took some good thinking and guessing on his part. Figuring out what aspects of improvement would be most important for a patient type when you don't really know what the patient is suffering from is a rather complex task which is not easily statisticized. In what way does your statistical program differ from factor analysis, profile analysis, etc., in this respect? I assume that once it is on the computer it will be untouched by human thought and be as stupid and rigid as the others.

The DDTD. Okay, I do have an answer for that one, but I don't know how good it is. First, let me finish off the procedure for condensing the types. We have regression equations, one for each treatment. If any two regression equations are identical then a subject will get the same predicted treatment score no matter which of the two treatments are being considered. Therefore, all regression equations should be tested for homogeneity and those which do

² C. J. Klett presented an example at this conference of a technique which uses the treatment-by-type interaction in the way suggested above. His work clearly anticipates my programmatic speculations.

³ Mimeographed version circulated before the conference.

not differ significantly should be combined. The number of distinct regressions corresponds to the number of distinct treatment types. Each regression equation can be used to allocate patients to its type.

Now, on the criterion side. Since we don't know which improvement score to use, let's use them all. We will put all the criteria into one battery, all the predictors into another battery, and do a Hotelling two-battery analysis. This will give us that linear combination of criteria that is most predictable from our predictors.

The C. Baby, that's a good answer—but to the wrong question. First of all, the Hotelling solution generally gives several linear combinations, not one. If you use the most predictable criterion, what do you do with the others? Secondly, maximizing predictability within each treatment group is not necessarily the same as maximizing treatment-by-type interaction. Thirdly, the Hotelling solution has to be found for each treatment group separately: but that means a different linear combination of improvement scores will be obtained for each treatment. How are you going to pick the highest of the k treatment scores for a subject when the metric for each treatment improvement score is different? How can you do a treatment-by-type interaction analysis when you shift to a different metric for each treatment?

The DDTD. Yes; I agree that Hotelling analysis via canonical analysis maximizes predictability and not interaction deviance, but I feel intuitively that it can not be too far from the optimal solution. On the other hand, one could work out the one linear combination of improvement scores that would maximize the interaction deviance. It should be a fairly simple exercise in matrix differentiation. But I can't help feeling that we probably should not be using the same improvement criterion for every patient classification. Unless some improvement rating can be given which most psychiatrists agree would have the

same meaning for depression, anxiety, hostility, apathy, etc., I would opt for a unique criterion for each type. This would place me in a quandary with respect to the treatment-by-type interaction. There would be a different interaction for every criterion.

The C. Yeah, well, *bon appétit*.

In this conference discussion session, Dr. Lubin made these additional remarks:

The Katz paper is a bit different because he tested the interaction of his patient types by treatment. He is rather casual about the whole thing, but I regard this as a very fundamental step. If we find that no matter which patient type you are treating, the same drugs have the same effect, (that is, if chlorpromazine is best for schizophrenics it is also best for depressives and it is also best for anxiety neurotics) then there is not much in the way of treatment indication that this new typology can give to the psychiatrist.

To the extent there is an interaction between the typology and the treatments, patients can be assigned to treatments so there is the best chance of recovery. In particular, the interaction has to be of the disordinal type, that is, the rank order of the drugs within each patient type must change. If the ranked order remains the same, then the best drug for any one patient type is the best drug for any other type of patient. It doesn't really matter.

This is a criterion for validating a typology. If the typology has a significant treatment-by-type interaction, it is of some use. It can tell the physician which drug is best to give for which group.

I would like to make one final recommendation. Let's junk the entire procedure of building up new typologies by cluster analysis, factor analysis, Q correlations, etc. There is no apriori guarantee that the typologies you get from the computer are any better than the typologies you could get by spinning dice and so on. At no time have I heard it suggested that a clinician could be used to generate new

typologies, a procedure which I think would have a lot more success than any computer program.

Discussant's Remarks

SAMUEL W. GREENHOUSE, Ph. D., *Chief, Theoretical Statistics and Mathematics Section, Office of Biometry, National Institute of Mental Health.*

First, I should like to make a general comment on the techniques that have been utilized by all four investigators in these papers, plus something which would tie into this morning's session, which was not mentioned but I believe is quite important. You will notice that all four papers—I believe Dr. Lubin emphasized this—made some use of factor analysis. In this morning's sessions, they referred to clustering, to discriminant functions and other techniques that were based upon a model which assumed linearity of relationships among characteristics.

As I understand it, many research psychiatrists and research clinical psychologists have stated that the problem they find most difficult in their research is to analyze those things which apparently are not linearly related. I believe this is very common, and it may also be inherent in the subject matter in this field. In this regard then it seems to me that the utilization of these techniques may be proceeding on wrong assumptions, on wrong bases. At the NIH several years ago I heard Dr. Guttman talk about the investigations he is making into developing techniques of analyses which would take into account nonlinearity of relationships. I understand that there are others working on this same problem. It seems to me that, to the extent variables studied in the behavioral sciences are related nonlinearly, it is most important that methodologists begin working on the development of valid techniques and proce-

dures to analyze properly the research data involving such variables.

I should also like to point out that in the Overall-Hollister paper they utilized a Bayesian approach to classification. I referred to this in my paper this morning, that there were statisticians who were beginning to emphasize Bayesian techniques in statistical analysis. I am sure you will be hearing more of this in the psychological and psychiatric journals in the future.

There is a point I would like to make with respect to Dr. Lubin's comments, which were written, but which he did not evidently have the time to make orally. He criticized rather severely, I suspect, a procedure in Dr. Katz's paper. In fact, I would like to quote Dr. Lubin on this: "The chlorpromazine treatment group was dropped from the analysis because of the small number of cases and heterogeneity of variance. I think this is a wholly unnecessary ritual sacrifice to a methodological rigor. Here and later in the use of the 'unweighted means' anova, a great many subjects were wasted. When the cost of each subject is considered in terms of money, energy, and time, it is not just bad statistics to discard available information, it is a sin!"

That was Dr. Lubin talking. Now, it seems to me it is common knowledge that you don't generally discard observations. However, here again just as in previous discussions at this conference, I think this must be considered within the context of the problem in which it occurs.

Clearly if an investigator is trying to determine the characteristics of a population of schizophrenics, say, and he wants to estimate certain properties of these characteristics, such as the mean level, or the extent of dispersion, discarding observations may result in a serious bias. If you obtain 10 observations, 1 of which turns out to be large, you obviously should not

discard it. In such a situation, all the classical principles would apply in all rigidity.

But note in this particular case Dr. Katz was considering comparisons. He was evaluating effects and responses to fluphenazine, thioridazine, or chlorpromazine and the placebo. All four treatments were applied to independent groups of individuals, and his primary interest was an assessment of the drugs against the placebo. A priori there were reasons for discarding chlorpromazine in the analysis, and these were not all due to small numbers and heterogeneity of variances. I believe Dr. Katz will refer to these a priori considerations.

But I think there is another problem here which I would like to bring before you, and that is this: If Dr. Katz was interested in these two drugs, A and B, and the placebo, and had done the experiment just including those three treatment groups, everything would have been all right. Nobody would have questioned it. Merely because he included a third drug group does not mean that Dr. Katz should be penalized in considering the relationships between the two other drugs and the placebo group. I think there are some fetishes built up about the discarding of information, and I think that it is not exactly as bad as Dr. Lubin implies. When the object is to compare drug A with placebo, discarding observations on drug A or on placebo may be bad statistics. However, discarding another treatment group need not be.

I also have one other comment. Unfortunately, it relates to the discussion of Dr. Lehmann, which you haven't heard yet, but we have had his paper and I did read it. On the initial program I think I was to follow Dr. Lehmann. The comment is as follows:

You recall yesterday afternoon someone made the comment he was happy to see the clinicians and statisticians snapping at

each other, and hoped this would continue. I, sorry to say, do not think that this should continue. In fact, I think this is very bad. It always happens at conferences where methodologists and clinicians get together, and I think it should come to an end.

What interests me is: What is the issue in this instance? I don't think that this has been clearly stated. Is it that the clinician believes his method of classification is the only valid one, and if so, for what? Is it for an understanding of the diseases? Is it for the purposes of therapy, for administrative reasons, or is it for getting a new building with additional wards? To me these are all appropriate and valid reasons. But we must consider the case where another investigator requires some classification criterion for some different purpose. Usually, initially he will make use of that system which is already in use. If, however, over time he finds that this system turns out to be defective, will the clinician deny the need for a new system to serve that end? I make this statement on the assumption that it is agreed here by everybody that there is a need for a classification system in the behavioral sciences and in the mental disorders. I make that assumption. I hope that it is correct.

Certain of the clinicians may ask how and why was the conventional procedure defective. I suspect it is at this point that the snapping begins. Whereas the clinician does have the right to ask in what way the procedure he has been using is defective, I think it is incumbent upon the methodologists or the psychologists, clinical or experimental, to respond. It is at this point, I believe, both groups should get together and explore the extent to which the purported defects are indeed defects. It is only by such joint efforts that we will make progress in developing a system of classification which will be of use to the greatest number of research workers and clinicians.

Discussant's Remarks

HEINZ E. LEHMANN, M.D., *Clinical Director, Douglas Hospital, Verdun; and Professor of Psychiatry, McGill University, Montreal, Canada.*

Having been invited to discuss the foregoing four papers in my capacity as a clinical psychiatrist—at least this is the role which I feel I should assume—I started with a question to myself: “What can a clinician contribute to a discussion of classification and typology?”

It is immediately clear what he cannot do in such a situation: He cannot constructively criticize or even very meaningfully discuss the methods of the statisticians. A clinician is obviously beyond his depth when it comes to an analysis of the comparative merits of multiple-discriminant functions, profile correlation indices, powered vector factor analysis, and the like.

However, there are two problems relevant to classification which a clinician may well discuss because they are in his domain—one entirely and one partly. Entirely within his area of jurisdiction, so to speak, is the manner in which the raw material for the statistical classifications and typologies is produced if it is derived from rating scales and questionnaires, because this brings clinical judgment back through a side door into an area which was supposed to have been cleared of it and its shortcomings.

The other problem which is only partly within the area of the clinician's concern deals with the end result of all classification, that is, the diagnostic types, the clinical syndromes, and behavior patterns which have been established—more specifically, their usefulness and their pragmatic value.

As I see it, the distinctive quality of the clinician lies in the fact that he is primarily neither an objective, detached observer like the physical scientist, nor a participant observer like the behavioral scientist, but rather a person who enters into a given situation for the avowed purpose of manipulating it and altering the

situation according to a subjectively conceived value-based plan. The clinician's first loyalty, as far as his general orientation is concerned will always be toward the empirical and the pragmatic. He is concerned with therapy and to that extent with a practical application of methods and procedures, but he is not concerned with methodology, i.e., with a theory of methods.

This does not mean that the good clinician is entirely at the mercy of casual trial and error, that he is dependent on guesswork or the kind of inexplicable intuition which moves the water dowser. Rather, a good clinician's approach should be solidly founded on scientifically established facts. He is allowed to disregard such facts in his activity only if they do not contribute to his success in manipulating the clinical situation, and he is allowed to extrapolate beyond these facts only if such extensions beyond scientifically established evidence will help him with his clinical manipulations. Furthermore, his informed clinical guessing or his clinical intuition which formed the basis of his extrapolations must be disciplined, consistent, and the considered result of systematic observations, even if they may still lack scientific confirmation.

Many psychiatric clinicians—and particularly those who take their work most seriously—maintain a certain skeptical and guarded attitude toward new systems of diagnostic classification. Why this conservative, if not reactionary, stand? Are psychiatric clinicians unduly pessimistic regarding the possibility of generalizing human behavior? Are they perhaps unnecessarily fussy and scientifically naive to a degree that prevents them from discerning major basic principles? Do they have vested interests—traditionally, prestigewise, even economically—in the individual approach to their patients?

Maybe one or the other or all of these reasons apply to some extent, but we must not forget that clinical psychiatrists have, after all, been the first ones to become preoccupied with nosological symptom-clusters,

syndromes, and diagnostic subgroups. All of our modern and sophisticated efforts at building new systems for grouping psychiatric patients, are they not the direct outcome of Kraepelin's descriptive *chef d'oeuvre* which, simultaneously with Freud's psychodynamic and Pavlov's psychophysiologic discoveries ushered in the scientific era of psychiatry at the turn of the 20th century? Phenomena, theories, and experiments respectively are the essential components of the achievements of these three giants. In some way it seems, however, that we are on more solid ground in psychiatric theory and experiment than we are in psychiatric diagnosis. This may be due to the fact that we are better able to control the crucial factors in theory and experiment, both of which we, ourselves, initiate—and that we are not capable of controlling the clinical phenomena which are given to us as end results of nature's experiments.

However, I stated already that the good clinician will have to base his general approach on scientific findings. There is no science without system and there is no system without an orderly arrangement of the factors involved. It is obvious that on grounds of parsimony one has to attempt to reduce the number of pertinent factors as much as possible. If there are many concretely given factors such a reduction is performed through the operation of clustering and classifying the factors. The physical scientist classifies his universe into a fixed number of elements and basic forces, the biologist has his classes and species, the psychologist works with cognitive, affective, and perceptual modalities, with fundamental drives and stimulus-response contingencies, and the physician reduces his problems by making diagnoses.

Now, the more complex the units are which one sets out to classify, the more factors will have to be matched and—unless one decides on arbitrarily restricted simplifications—this must, at times, result in very small classes. If the number of criterion factors is zero, the class will comprise the

whole population of the available sample. If the number of discriminating factors is infinite, then each individual in the sample becomes his own unitary class. Between zero and infinite the number of criterion factors determines the size of a class or homogeneous group. Since human beings are the most complex subjects extant it follows that the number of important criterion factors is likely to be high with humans and the populations of each sample group will tend to be correspondingly small, unless the original pool of available unclassified subjects is large.

One problem that the clinician probably exaggerates and the statistician takes too lightly is presented by those patients showing clear-cut psychopathology who cannot be properly classified. The clinician makes much of these unmatchable idiosyncratic cases who are—each single one—in a class by themselves. Lorr's two examiners agreed that 34 percent of the male patients they had been rating were unclassifiable. By establishing cutting scores and certain classification rules and applying them sequentially by computer to all cases in his sample and by using correlations of syndrome profiles, Lorr could eventually classify 80 percent of all cases and he thinks that by the use of multiple-discriminant functions the 20 percent of unclassified cases could still be further reduced.

Overall and Hollister state the problem of patient classification as one of requiring enough detail to permit representation of important individual differences, yet retaining enough parsimony to conceptualize easily the classification prototypes. But how much detail is enough and how much parsimony is adequate? The nicety of judgment which is required here cannot be supported by scientifically derived rules of decision, but resembles the kind of judgment clinicians are usually called upon to make.

At this point I should like to review what are generally held to be the most convincing reasons for developing classifications of

pathological conditions. First, diagnostic types are required as basic material for statistical, biometric, and epidemiological studies. Second, they are necessary for research into the etiology of diseases. Third, a diagnostic classification can often serve as an indicator for the specific therapy which needs to be applied. Fourth, diagnostic groupings frequently enable the clinician to predict with considerable precision the most likely future course the disease will take. The first two reasons are of particular interest for the researcher, the last two for the clinician.

Now, the clinician is guided by pragmatic considerations and if the classifications which he recognizes on the basis of tradition, experience, and nonconfirmed notions allow him to treat his patients successfully and to be fairly accurate in his prognoses, nobody will cast aspersions on the unscientific pedigree of his diagnostic classes—at least, not as long as the clinician can be explicit about his diagnostic criteria so that he can teach others the rules of his diagnostic craft.

The researcher, however—it has always seemed to me—should establish some firm and objectively verified criteria for his clusters and classes. This means he should be independent of the clinician's subjective judgment. Since mental disease manifests itself mainly in behavioral deviations, the behavioral symptoms serving as the criterion factors for psychiatric diagnoses should then be capable of objective and unambiguous recording. This is, of course, only possible for a very limited number of behavioral manifestations; namely, for those that are capable of physical registration, for instance, the amount, speed, and direction of a patient's movements, the amount and intensity of his vocal productions, the number of specific sounds or words he utters in a given time, etc. But most of the important behavior items which have to be recorded by the research scientist who searches for new and better methods of diagnostic classification are of the type that have to be

judged by human observers and have to be recorded in rating scales, usually in some measure of intensity which depends on the subjective estimate of the rater.

What has to be rated are such deviations as conceptual disorganization, hostile belligerence, anxious intropunitiveness, depressive mood, anxiety, guilt feelings, happiness, conformity, somatic concern, and similar complex phenomena. It is true that many of these concepts are broken down in the rating scales into a number of simpler behavioral manifestations, but no matter how much one simplifies the rater's task it will, in most cases, still remain a subjective recording of human phenomena which have a distinct gestalt quality which can only be appreciated by another human being. This—at least in my opinion—transposes the behavioral rating from the objective scientific dimension into the subjective clinical one.

One may argue here that a modern well-constructed behavior rating scale has broken down the task of observation into such simple steps that anyone with some sense and a certain amount of training can learn to perform as well—at least as reliably—as some piece of physical apparatus. To this I would reply that the reliability between various observers commonly accepted as the next best thing to validity is often not more than an artificially produced conformity of well indoctrinated and over-trained subjects. As for the validity of behavioral ratings, particularly when it involves the evaluation of feeling states and attitudes, I do not believe that one is allowed to assume that every normal person either possesses or can acquire the ability to judge accurately such behavioral manifestations or emotional states. Some experiments conducted at the Verdun Protestant Hospital on which we have recently reported have indicated that certain persons seem to be able to make reliable and highly valid judgments of the behavior of psychiatric patients without ever having received any systematic training or having had much

experience in this field, while some experienced professionals in this field after many years of training fail to come even close to the performance of these naive, but somehow gifted observers.

In other words, while there is undoubtedly a high correlation between clinical training/experience and the validity of psychiatric observations which are to be recorded in behavior rating scales, there is also another, apparently primary, factor involved which is entirely independent—should I say orthogonal to experience and training? In some observers, this peculiar factor seems to be the most important one in accounting for the validity of their behavioral observations.

It would appear then that in making the observations which are required for the behavioral rating scale record, the research scientist may be at the mercy of certain factors which, at the present time, are uncontrollable. If this is so, he is working with the same tools as the clinician and we have completed the circle.

A radically new approach to the problem of classification would have to avoid all clinical judgments, whether they are global or fragmented into computer-sized bits. As an example, a really new classification of heart diseases became possible only after the electrocardiograph had been invented and had provided data on cardiac function which could not have been obtained in any other manner. Yet, the clinical methods of auscultation and percussion had been employed for almost a century and had served the clinician well. If one wishes to measure the size of the heart, auscultation and percussion are still useful clinical methods, although they are notoriously imprecise and are replaced by X-rays whenever the facilities are available. Pursuing this comparison one might say that the clinical evaluation of behavior is as serviceable and as precise as the best possible results of auscultation and percussion are in cardiology, but it would be highly desirable to have an objective method—an X-ray of behavior—to assure

consistently valid findings by anyone taught to apply such a hypothetical objective technique to behavior.

Unfortunately, there is, so far, no method known that would allow us to obtain objective and reliable data which have a valid correlation with complex behavior variables, and are independent of the judgment of human observers. Some experimental work that we have recently completed involved the systematic testing of more than 200 psychotic patients with a battery of 14 psychophysical test procedures. To our disappointment we found that specific diagnostic test patterns could be established through factor analysis only in the chronic psychotics who had been hospitalized for many years, while no distinct diagnostic test patterns could be isolated in the sample of acute psychotic patients. Their performance on the 14 tests yielding some 20 different measures with loadings on 4 different factors could only be arranged on a continuum from better to worse, regardless whether the psychoses to be classified had clinically been diagnosed as manic-depressive psychosis, senile psychosis, schizophrenia, or psychosis with psychopathic personality. All these diagnostic categories were flattened into a unidimensional continuum of psychophysical deficit. Some well-preserved patients suffering from senile dementia and schizophrenia scored higher than the less preserved manic and depressed patients and, of course, also very much higher than other more severely demented senile psychotics and schizophrenics. It seems that in the acutely psychotic patient there is so much psychophysical noise that the central information is completely obscured or rather reduced to a unidimensional scale of total personality functioning. This noise which is possibly the reflection of emotional reactions of the still preserved personality to the psychotic process, disappears only after years of hospitalization and, at that time, some specific diagnostic test patterns begin to crystallize out.

We have convinced ourselves that even a sophisticated analysis of the data obtained from the performances of psychotic patients on a representative variety of the best known psychophysical tests does not yield diagnostic patterns and classes which are as rich and useful as those which clinicians have already had in their possession for many years.

We still hope, however, that new and ingenious methods will be developed which will eventually allow us to classify mental diseases on a purely objective basis. Some of the recently developed electrophysiological techniques or even some of the—up to now—unproductive psychophysical and psychometric tests in combination with psychopharmacological loads may hold the answer.

While I do not believe that the classification methods based on behavior rating scales, inventories, questionnaires, and structured interviewing schedules have any true heuristic value that is greater than that of other entirely clinical methods, I must, of course, admit that the methodological application of clinical judgment which is enforced by these diagnostic instruments has considerable advantages over the unstructured clinical approach when it comes to tasks which involve the comparison of clinical data obtained from different persons at different times as in epidemiological, in biometric, and in comparative therapeutic studies. They are also very important for the teaching of clinical techniques and skills and finally, serve the purpose of systematic reviewing and stocktaking of the clinician's available tools.

But there are three particular criticisms which I should like to express regarding the cluster analysis of the rating-scale-derived items. They are:

1. No attempt is made to establish the relevance of all the symptoms and phenomena that are considered in the statistical search for the assumed diagnostic pattern. It has, of course, been pointed out that one cannot establish such relevance without hav-

ing first considered all factors on an equal basis if one does not want to prejudge the issue. But, such factors as anxiety, hostility, and excitement which all exist to some degree normally in every person, may carry very little, if any, specificity and to speak of the hostile paranoid and the quiet paranoid might not be any more or less meaningful than to distinguish between the agitated, the anxious, the demanding, the indolent, and the disciplined type of appendicitis patient. No doubt such characterizations may be extremely important for the patient's therapeutic management and his prognosis—but they do not contribute to a specific diagnosis.

2. The hierarchy of psychotic symptoms is disregarded by most cluster-analytic methods—again for reasons of scientific neutrality. To the clinician, however, such neutrality appears to be almost contrived, when, for instance, hallucinations, delusions, lack of self-care, posturing, verbigeration, and mirror gazing are all treated as though they were of equal significance for the diagnosis of schizophrenia, although the first three of these symptoms may occur in just about any psychotic illness, while the last three—in combination and sometimes even singly—are almost pathognomic for schizophrenia.

3. The interaction and interdependence of different symptoms cannot be taken into account in most of the cluster-analytic approaches. It may not be true that most patients with delusions of persecution will also sooner or later develop delusions of grandeur and vice versa as some clinicians believe, but there is little doubt that such interdependence does exist in a number of paranoid cases. The hostile and the grandiose paranoid may, therefore, only represent two different stages of the same disease process. There are many other examples of such interrelatedness of different symptoms and for the clinician it is extremely important to have at least some working hypotheses in this area. To mention only one other example: Some depressed patients may

become extremely anxious and panicky because they feel threatened by the emotional and ideational emptiness which they experience as an essential component of their depressed state. Conversely, other depressed patients have become depressed as a result of exhaustion from excessive and sustained anxiety.

Two of the principal arguments which are advanced to justify the search for new diagnostic classifications in mental diseases are:

1. That certain particular symptom patterns might respond better than others to specific forms of treatment and especially to specific forms of pharmacotherapy.

2. That certain symptom patterns might be associated with a specific antecedent factor in the history and a specific prognosis so that the clinician would be in a better position to prevent and predict the development of such diseases.

This is quite true, but it should be remembered that many of the traditional diagnostic classifications of mental illness also enable the competent clinician to successfully choose the best possible treatment and to predict—often with considerable accuracy—the course and final prognosis of the illness. True, this is valid only for those psychiatric clinicians who are giving careful consideration not only to the gross, but also to the more subtle aspects of psychopathology and who for that reason might be called experts.

However, any data derived from behavior rating scales and inventories also depend on the clinical ability of the rater as has been shown and it is questionable whether the higher reliability of the data obtained through the employment of such diagnostic instruments has any true clinical significance. Or should we conceive of such instruments as the poor man's clinical acumen, some sort of instant clinical competence in packaged form?

Having taken such an emphatically critical stand in regard to statistical methods of diagnostic classification I shall now, for a

while, abandon this attitude of the devil's advocate which, I must confess, did come quite naturally to me much of the time and I should like to discuss what I consider to be the most essential issues brought out by the four authors whose papers have been presented this afternoon.

What strikes me as most important in the paper by Grinker and Nunnally is their basic approach to the clinical problem of depression. They are convinced that a new sharp look at the clinician's methods of dealing with depression is required. However, they do not set out to create a new clinical entity *de novo*, so to speak, from an amorphous mass of factors, but accept the syndrome of depression as a preformed and clinically given entity. Their scientific detachment displayed in a systematic review of symptoms does not extend to a complete antiseptic abolition of such clinical tools as intuition and empathy. It is detachment only from the adherence to time-honored stereotypes, traditional biases, and theoretical ballast resulting in true observation and careful description of clinical phenomena.

The authors found that silent intuitive processes of raters resulting in judgments on the patient's feelings and concerns were productive of more agreement among the psychiatrists than their observations on the patient's overt behavior. In other words, the psychiatrist's inferences about patient's private emotional processes are more reliable than his observations of patient's public performances. This does not inspire particular confidence in behavior rating scales which are the products of psychiatric observers.

As for the validity of ratings, Grinker and Nunnally could show that the amount of information available to the rater is more important here than the rater's professional standard.

The authors factor analyzed the behavioral patterns of traits and those descriptive of feelings and concerns and arrived at 5 factors of the latter and 10 factors repre-

sentative of depressive behavior. Commenting on the small number of feeling-concern factors which resulted from this factor analysis, Grinker points out that there is, after all, only a limited number of final common pathways which can serve as an expression of deranged mental and emotional processes. Interestingly, the patterns descriptive of the patient's feelings and concerns did not correlate with the patient's current behavior patterns which is not too surprising in view of the different reliability of psychiatric observations in these two different dimensions.

From a combination of the 5 feeling and concern factors and the 10 behavioral factors Grinker and Nunnally derived 4 factor patterns representative of 4 depressive syndromes to which they refer with admirable scientific restraint as factor patterns, A, B, C, and D. I have not been able to avoid giving simple names to these four types and prefer to refer to them as the empty, the anxious, the hypochondriacal, and the angry type of depression. One wonders by what strange inversion the clinician employing statistical methods has refrained from the clinical conceptualization of the classification end product, while the statistician using clinical material usually ends up with labels on his types which are borrowed from psychiatric terminology, although it is always quickly pointed out that these terms are not to be understood as meaning the same thing as in the current diagnostic language. Is the clinician leaning over backwards to remain scientific and the scientist straining hard to appear clinical?

I should like to mention only one of a number of minor points which I would question in this paper which, to me, seems to be a fine example of constructive scientific progress applied to the clinical methods. And that is the authors' statement that if there are equally extreme symptoms present in the areas of melancholia and disruptive thought processes, then the diagnostic classification must be made in terms of the more dangerous symptoms and it

would necessarily have to be schizophrenia. I cannot accept at face value the implication that disruptive thought processes are more dangerous than symptoms of melancholia, nor do I see the logical need of calling such a patient a depressed schizophrenic, rather than a depressed patient with schizophreniform symptoms. Where is the evidence that depressive symptoms can enter into schizophrenia, but that disruptive thought processes cannot enter into depression? There is, in my opinion, no clear-cut rule for decision in such ambiguous cases—the final diagnosis would have to depend on the clinician's careful weighing of all the information available.

In contrast to these workers, who presuppose the existence of a depressive state on the grounds of clinical evidence before they subject this condition to statistical analysis, Lorr approaches the whole field of the functional psychoses like a *tabula rasa*. One might ask here how he was able to establish the existence of functional psychoses as such? A radical nonclinical approach to the classification of mental disorder would really have to start from scratch and distinguish through some systematic procedure the functional psychoses from the organic, from drug-induced psychotic states, from the psychoneuroses, from the behavior disorders, from hypnotic trance, and from malingering, to name only some of the conditions which might produce a pathological score on the rating scales. I can almost hear the angry exclamation of some in the audience: "For heaven's sake, be reasonable! After all, can't you accept the psychiatrist's judgment at least on the issue whether or not somebody is psychotic?" Indeed, I can, but having accepted that much I also can accept a great deal more as far as the psychiatrist's competence in diagnosis is concerned.

Lorr has established 10 basic psychotic syndromes and the profiles of his pathological types are characterized in terms of mean syndrome scores that lie above or below the general mean on a standard score

scale. His comprehensive and well-tested measuring instrument is the inpatient multidimensional psychiatric scale (IMPS) which is now one of the most widely used scales for the scientific study of psychotic patients on the North American continent. The scale characterizes the various clinical types by digits which refer to the 10 psychotic basic syndromes. His anxious-disorganized type, for instance, is characterized by the digits 3, 5, 6, 7, 8, 9, and 10. While such characterization by digits is conceptually meaningless, it is, nevertheless, a valid scientific designation indicative of a number of essential attributes. It might, in fact, be better to leave the naming of the types to each individual clinician who would then at least know what he means by his own shorthand description. For convenient reference in scientific discussions the different types could be labeled with letters of the alphabet.

In a sample of acute male psychotics Lorr found nine types in four kinds of patient classes. In a sample of acute female psychotics he also found nine, but two were different from the male. Chronic male patients yielded seven types and in this chronic group the only new subgroup was the retarded, while the excited and hallucinated paranoid types did not appear in the chronic sample. In long-term patients seven subgroups appeared through cluster-analysis. Lorr points out that a sample of carefully diagnosed schizophrenics includes nearly all psychotic types found among a random sample. In revalidation studies performed on different samples, certain differences in the distribution of patient types were noted which Lorr explains on the grounds of different sampling, differences in rating procedures, and rater competence, and different exposure to drugs.

In the author's system there are four paranoid groups: The excited, the hostile, the grandiose, and the hallucinated. He mentions that the excited paranoid resembles the classical manic type and one wonders whether he means that these are

manic-paranoids or paranoid-manics? Since Lorr does not presuppose the existence of Kraepelinian diagnostic entities, I would assume that he simply wishes to remind the clinician that the excited paranoid is what many psychiatrists would call the manic type. The question, however, remains whether it is more justified to speak of an excited paranoid than it would be to speak of a paranoid excited, in which case the excitement would be the pivot trait and there would be several groups of excited patients instead of four types of paranoids.

In Lorr's system the depressive state is reflected in the intro-punitive and the retarded types. The validating evidence of one of these types is represented in the fact that intro-punitiveness is very likely to be diagnosed as depressive reaction, that such patients receive psychotherapy more often than any other subgroup and that they tend to be discharged within 6 months after admission more often than most other patients. Members of the intro-punitive type are also 92 percent white, likely to be married and to come from upper class urban areas. While accepting these facts as characteristics of depressed patients I wonder how many psychiatrists might wince at the label intro-punitive which, at best, leaves one with the unwarranted assumption that all dysphoric patients are punishing themselves and at worst, conjures up the hypothetical construct of magical introjection which often has considerable pragmatic value in the one-to-one psychotherapeutic situation, but is difficult to fit into an austere presentation of computer-derived behavioral factors.

Personally, I am impressed by the author's superb workmanship, somewhat confused about the meaning of his complex system and not at all convinced that it is useful for the clinician. His method has proved to be most valuable as a means of comparing therapeutic results flowing from widely dispersed sources into one great pool for systematic evaluation, for instance, in NIMH-sponsored collaborative studies on the action of drugs in schizo-

phrenia. These studies were necessary, definitive, and so complex in their logistic requirements that orderliness and comparability of data counted more than their individual meaningfulness.

Overall and Hollister expose three major problems which account for the poor reliability and validity of psychiatric observations:

(1) Vague and disorganized verbal characterizations account for the lack of consensual understanding of the distinctions between the various diagnostic types.

(2) There are considerable difficulties in evaluating reliably psychiatric characteristics.

(3) There is a lack of clear-cut decision rules for combining multiple observations—"• • • The standard nomenclature is completely silent on this important point • • •."

The authors go on to state that lack of normality in distribution of the data is one of the worst problems. In one attempt to test the consensual validity of the psychiatric diagnostic nomenclature, they asked 38 experts to construct 13 typical patients of 13 functional psychotic types on their brief psychiatric rating scale (BPRS).

Through factor analysis of their rating-scale data the authors have established four orthogonal higher order factors of psychosis: Thinking disturbance; psychomotor disturbance; paranoid interpersonal disturbance; depressive disturbance. They, furthermore, identified five modal profiles of schizophrenic patients and three modal types of depressives, pointing out that these eight modal types are difficult to conceptualize.

Although they found anxiety in both schizophrenics and depressives, each clinical type has one modal type which lacks anxiety. Hallucinated behavior is confined to a single modal type. Hostility is found in one schizophrenic and one depressed type. Depression is low in all schizophrenic modal types.

In addition to their four higher order factors and their eight modal types they also distinguish three major symptom clusters: The paranoid, the schizophrenic, and the depressive. It is explained, however, that these are not to be considered as clinical diagnoses, but only as symptom profiles derived from patients rated on the 16 variables on the BPRS and most resembling these 3 prototypes.

In one intriguing statement the authors reassure us that while computers may be necessary in the research phase, it might well be that they could later be replaced by simple stencils for clinical applications.

The most promising applications the authors have made of their own elaborate factor-analytic method lies in their systematic testing of patients' treatment responses in relation to their pretreatment symptoms and—more importantly—in relation to specific drugs and classes of drugs. They found, for instance, that of their three depressive profile types, the anxious depressive patients responded better to thioridazine and the retarded to imipramine. Among schizophrenic patients they observed that phenothiazines and triperidol were more effective in paranoids, and oxypertine and imipramine gave better results in the depressed type profiles.

Finally, they propose four types of patients as being differentially responsive to four types of drugs. One type of patients responds best to acetophenazine, perphenazine, and oxypertine, another to benzquinamide and thioridazine, a third to trifluoperidol, and a fourth to imipramine.

The authors suggest that psychiatric typology may be based on drug responses as a technique for identifying patient populations and they feel that the practical value of this could amount to improvement of as much as 50 percent in overall therapeutic efficacy of psychiatric drug treatment.

No doubt the authors have a lot to say that is of interest for the clinician. One must still question, however, whether the

differences in therapeutic response which they found are great enough to be of practical significance. It could be that the differential therapeutic responses are of such small magnitude that they might disappear with a change of doses of the two drugs or after waiting 1 or 2 weeks beyond the final rating. While the authors have been able to devise a method which allows them to compare the therapeutic responsiveness of patients who were studied at different places and by different observers with widely differing degrees of competence, they have not established evidence that one or two competent clinicians working with a much smaller sample of patients and not employing sophisticated statistical methods could not have arrived at similar conclusions. That the authors have resorted to the construction of clinical stereotypes by expert clinicians in order to validate their computer-derived patient types suggests that they themselves still have considerable belief in the criterion qualities of the clinician.

Katz begins his paper with a discussion of the problem of validity in patient typology and selects consensual and pragmatic validity as the two most important types in this area. Explaining that typological theory in the personality field is derived from gestalt principles he goes on to say that this holistic way of viewing the organism appears to be congenial to clinical thinking. He then outlines the research steps of his own project which are:

- (1) The selection of the significant variables in the individual's pattern of social behavior.

- (2) The development of a method for quantifying the various facets of the individual's social functioning.

- (3) The determination of configurations of behavior manifestations.

- (4) The consensual validation of the typology.

- (5) The predictive validation of the typology.

Katz has noted that a major gap in research of psychotic disorders is our lack of information on the patient's behavior prior to entering the hospital. To furnish this he turned to the patients' close relatives who were interviewed using a specially constructed rating inventory on which the relatives' observations of the patients' early disturbed behavior in the home were recorded. A profile of the systematic and social behavior was then developed through a factor-analytic method. In order to determine whether the patient types described by the relatives could be located within a heterogeneous population of patients classified as schizophrenics the *Q*-procedure of factor analysis was applied to a matrix of intercorrelations of the rating profiles of 24 acute schizophrenics.

At this point, the author discusses the problem of selecting a particular technique for assigning subjects to clusters and categories and points out that such a selection would have to be based partly on statistical considerations, partly on the nature of the method of measurement being used, and partly on the content of the data. His own choice of method was based on his judgment that given an observational rating procedure and considering the fact that each patient was rated by a different observer, configurations represent the most relevant and reliable information with which to describe patients.

Six patient types were found to account for most of the variance in the matrix and these patterns were found to be meaningful from a clinical standpoint in the sense of being recognizable types of patients as judged informally by clinicians.

Katz found three paranoid subtypes:

- I. An agitated, belligerent, fairly well-integrated type.
- II. A withdrawn helpless type.
- III. An expansive, agitated, less integrated type with bizarre ideation.

It appeared that at least four of the types which had been described by the relatives

were in agreement with the observations by other observers, namely, nurses and psychiatrists. Certain differences were noted, e.g., the psychiatric raters did not see the extreme anxiety in patients which had been seen by the relatives. It is not clear whether this difference is a function of differences in observing or due to the fact that the patients had lost much of their anxiety after entering the hospital.

Having thus established consensual validity for the patient types derived from the relatives' observations, Katz then examined the question of predictive validity of these types by analyzing the responses of schizophrenic patients to different drugs in collaborative studies sponsored by the NIMH. He found that the acute panic state responded best and the withdrawn, periodically agitated group least to all drugs used in this study. It was also observed that the belligerent paranoid does less well than the withdrawn paranoid and it was further indicated that fluphenazine was more effective than thioridazine in reducing schizophrenic disorganization in the withdrawn suspicious, while thioridazine was significantly more effective in the withdrawn, periodically agitated.

The difference in therapeutic responsiveness of particular patient types to particular drugs does not seem to be of an order that would carry much practical significance. The fact that certain acute and anxious patients respond better to treatment procedures than certain withdrawn types of schizophrenics is not new to experienced clinicians, and it would appear probable to me that a competent clinician would be able to make valid predictions of the differential therapeutic responsiveness of the withdrawn and helpless patient versus the withdrawn and periodically agitated one. Although this may appear to be a gratuitous statement after the fact, I would expand it even further to the proposition that a better therapeutic response of the withdrawn paranoid than that of the bel-

ligerent paranoid might be predictable on clinical grounds alone.

What is needed now is a controlled study of clinical prediction versus statistical prediction in the area of psychiatric therapeutics. There seems to be no other way to settle the question which method is more precise and more practical.

Certain findings in the tabulated data of the relatives' ratings of schizophrenics resulting in the description of six basic types are rather unexpected on clinical grounds. For instance, the withdrawn, helpless, suspicious type is rated low on anxiety and nervousness and the agitated, helpless is also low on anxiety. I would think that anxiety and nervousness in these patients might have been present to a considerable degree, but could not be detected by the limited number of questions probing these areas with the special inventory used. Katz recognizes the practical and methodological value of some clinical approaches when he uses the global improvement rating as his criterion for the patient's response to treatment in his statistical analysis of differential responses.

In my opinion, the most important achievement of this work lies in the demonstration that the information supplied by psychiatric patients' relatives can be used as reliable and valid material for further systematic analysis. This is a new and welcome addition to the clinicians' as well as the researchers' supply of observational data as is perhaps best expressed in one of the author's sentences: "• • • Despite their presumed differences in conceptual frameworks, despite the differences in measuring instruments used, and despite the variability introduced by using a large number of raters from nine different hospitals, the psychiatric rater's descriptions of the behavior of the various patient types are highly similar to the relatives' descriptions • • •."

If I may be permitted to make one other exasperated comment as a clinician looking at these overwhelming arrays of power-

ful factor-analytic efforts, all for the avowed purpose of helping the naive clinician to obtain a clearer view of his field, I should like to say that such efforts strike me sometimes like an attempt to establish the grammar of a language scientifically through factor analysis. A researcher trying to do this might find, to his amazement, that there is indeed a cluster of words that corresponds to the commonly accepted notion of nouns, another to the category everybody naively thinks of as verbs, a third cluster that looks almost exactly like the so-called adjectives, etc. If my comparison appears to reflect a rather insensitive view, then I would think that statisticians have an obligation to communicate more convincingly to the clinician the meaning and value of radical factor-analytic contributions in regard to the nature of well-established clinical syndromes.

If factor analysis will—as I believe—only get out of a mass of data in unscrambled form what has gone into it in scrambled form, then factor analysis would have to be considered as an effective decoding device with little or no heuristic power. Whether such decoding of complex data and the resulting detection of invariant features in the behavioral field could not be done as reliably, yet more rapidly and more validly—although not always as conveniently—by clinical methods, is a question which still has to be answered.

In summing up my own ideas which I tried to express here in a rather rambling fashion, I would say that I look at the statistical analysis of behavioral data—in particular, the factor-analytic method—as a convenient, neat, and systematic methodological device which has the advantage of allowing for automatization of the verification process, but has the disadvantages of being only partly objective and possessing little heuristic power.

(In the discussion session of the conference, Dr. Lehmann made these additional remarks:)

What strikes me as most important in the

paper by Grinker and Nunnally is their basic approach to the clinical problem of depression; namely, they did not take just an amorphous mass of data, but first made the clinical decision that there was such a thing as a depressive syndrome or state, and within this depressive state, which was clinically determined, they began to tidy up things and do away with stereotypes, and get new and probably quite fertile classifications.

They also found that silent, intuitive processes of the raters resulted in judgments on the patients' feelings and concerns which were more productive of agreement among the psychiatrists than their observations on the patients' overt behavior.

This is very important because most of our rating scales and most of the data that are fed into the computers, of course, depend mainly on observations of manifest behavior, but if that is less reliable than the psychiatrists' reports on the patients' feelings and concerns, it should make us think.

I was impressed by the fact that the scientific detachment of these workers did not extend to a rejection of tools such as intuition and empathy. These are perhaps not very scientific but from the clinical point of view still indispensable.

I feel that this was an excellent way of using the new techniques that we now have—cluster analysis or factor analysis and various computer techniques—to tidy things up within a clinically given entity and to rediscover some things that we had forgotten about or hadn't seen yet, and also to abolish stereotypes and prejudices.

Now, Lorr, on the other hand, approaches the whole field of the functional psychoses like a *tabula rasa*, as though we didn't even know that there are such things as psychoses. How do we know? I feel a truly radical nonclinical approach to the classification of mental disorders would have to start from scratch and distinguish through some systematic procedure the

functional psychoses from the organic, from drug-induced psychotic states, from the psychoneuroses, from the behavior disorders, from hypnotic trance states and from malingering, to name only some of the conditions which might produce a pathological score on the rating scales.

Personally, I was very much impressed by Lorr's tremendous workmanship of which I understand little, and I am somewhat confused as a clinician about the meaning of this complex system, because to me in my work with patients it just isn't functional and I am not at all convinced that it is very useful for other clinicians.

It is useful for research efforts, such as the collaborative studies of the NIMH, for instance, where one needs comparative data from a pool of unrelated hospitals. For specific research projects, sure enough, such scales are valuable, but otherwise the individual experienced clinician, I think, will hardly get very much from it.

Now, Overall and Hollister—well, it is, of course, very impressive what they have found; namely, as Dr. Lubin already elaborated, the drug response groups—various groups of patients that respond to different chemical or pharmacologic groups of drugs, and they also suggest that psychiatric typology may be based on such responses as a technique for identifying patient populations. They feel that the practical value of this could amount to improvement of as much as 50 percent in overall therapeutic efficacy of psychiatric drug treatment. No doubt, they have a lot to say that is of interest for the clinician.

Personally, I still question whether at this point there really is very much that these techniques which are elaborate, complicated, and take a lot of time, can tell us that clinicians couldn't find out for themselves with a much smaller sample in a much shorter time. Perhaps it won't always be so, and if in the future the computers will be able to overtake the clinician, fine. Perhaps they have already overtaken him, but I think this still ought to be shown.

Coming to Katz, I think the main value of his paper lies in the demonstration that the information supplied by psychiatric patients' relatives can be used as reliable and valid material for further systematic analysis. This is a new and welcome addition to the clinician's as well as the researcher's supply of observational data, as is perhaps best expressed in one of his own sentences: "Despite their presumed differences in conceptual framework, despite the differences in measuring instruments used, and despite the variability introduced by using a large number of raters from nine different hospitals, the psychiatric raters' descriptions of the behavior of the various patient types are highly similar to the relatives' descriptions."

So we have really gained a new source of data since we know now that relatives' information is reliable.

Since I used the word, may I make just a comment on reliability. Reliability of testers or examiners, in my opinion, is often—certainly not always, but often—only the result of overtraining and indoctrination by a strong, dominant personality, who is doing the training. In these cases reliability really has no intrinsic relation to validity. In other words, the same raters would not be reliable if they were rating at different places or different hospitals and not under the dominant influence of perhaps one or two persons who are responsible for them.

Perhaps I should answer Dr. Greenhouse's challenge. I don't want to snap at statisticians or biometricians or experimentalists. I simply want to express my impatience at their not having discovered, clever as they are, that the data they are getting just aren't worth all of the trouble they are going to. The clinicians don't give them any data that are sufficiently reliable. You know the beginning of the famous recipe for making a rabbit stew: First, get yourself a rabbit. Well, there may be a lot in the stew, but it may be cats and rats and all kinds of other organic matter and you can't be sure at all that there are

only rabbits in it. This brings me back to Dr. Grinker's very serious plea for more careful data gathering. Things are really in a bad way. I know, because I am constantly dealing with the problem at a university where we think we have very high teaching standards, and where it is almost impossible to come to any general agreement, even on the simple behavioral phenomena, such as what is bizarre, what is confused, what is depressed, what is apathetic, what is anxiety. What are guilt feelings, happiness, conformity, productivity, hostile behavior, belligerence? All this would take hours of seminar sessions and discussions and much training and demonstration to get the residents or the young psychiatrists to recognize these phenomena when they are there and not to misidentify them.

Now, very few are taking this trouble. Most of the money, too, is being put into the programs for evaluating what we really haven't got—and we have very ingenious methods for doing this, very refined methods. We have the computers where we can deal with—well, an almost infinite number of data, but we haven't any good methods of getting reliable data. In this respect we are far, far behind. There are, unfortunately, few clinicians whose data one could take and say, "All right, I believe it," and then, only then, let the computer go to work and make something of it.

Now, why this is so is another question. As Dr. Grinker pointed out, all of the emphasis has been on theory, psychodynamics, on understanding of the meaning of what is going on, and also on what to do with the data once we get them, and practically nobody has bothered how we get them. It is like a physicist who must read the indicators of his instruments, but has forgotten his glasses and just guesses at what the readings might be. Well, no matter how great his theory is and how many computers he has to work with, these raw data, they just aren't worth anything. Very few people have glasses on nowadays in

getting clinical psychiatric data, even the simplest.

Yesterday Dr. Kety was doing away, very casually, it seemed to me, with operational definitions. He said, "We have to have observations or definitions that have meaning." I think that is a fallacy. If we insist on that, then we immediately focus again on theory or on evaluation before we really have obtained precise data. In the meantime, while we are gathering data which we do need for clinical purposes; i.e., clinical classifications, let's make operational definitions. They are scientifically quite acceptable, and at least they will give us the reliable raw material which we might later make more meaningful.

There is one other thing I would like to point out, and that is the clinician's peculiar position. He is not a detached observer as a physical scientist is, and he is not a participant observer as a social scientist is. He is a manipulative observer. He observes only in order to change the data he is observing. Right from the start that will bias him as he will look for systems that will help him to do his manipulating better. But, if we look for comparative results, if we look for pathogenic mechanisms or neurophysiological or biochemical mechanisms, then we need data that are not empirically oriented towards manipulation. But let me repeat again, we need data. It is not too difficult to get precise data about certain behavioral manifestations, as for instance, the speed and direction of a patient's movements—they can be filmed and measured objectively—the amount and intensity of his vocal productions, the amount of specific sounds or words he utters in a given time. All of these are objective data, and so are, of course, also physiological and pharmacological and biochemical data. But for anything that involves emotional overtones—such as anxiety, depression, even of the simplest sort, or hostility—we do need a human instrument because there is just no other instrument extant that can measure such

things reliably in any way. So we have really an overblown evaluation of data for which we have not developed enough good instrumentation that will allow us to get them adequately. But we have not enough good human observer instruments.

In closing, I would say that if factor analysis will, as I believe, only get out of a mass of data in unscrambled form, what has gotten into it in scrambled form, then factor analysis would have to be considered as an effective decoding device with little heuristic power. Whether such decoding of complex data and the resulting detection of invariant features in the behavioral field could not be done as reliably, yet more rapidly and more validly—although not always as conveniently—by clinical methods, is a question which still has to be answered.

OPEN DISCUSSION

(At the conclusion of Dr. Lehman's remarks, the conference members were invited to participate in an open discussion, moderated alternately by Dr. Kelly and Dr. Lubin.)

Dr. ZUBIN. Dr. Katz's point about personality being the basic underpinning of his work is a very important step in the direction of introducing personality theory into this whole picture. Dr. Katz mentioned Jung's typology as providing the kinds of considerations that are important for considering the relation between personality and psychopathology. There are three possible ways of looking at this relationship: First, they are one and the same—a person is born with his psychopathological potential or vulnerability, or in his early life he has already developed it; the rest of his life is simply a flowering of his personality and of the psychopathology which corresponds to it. Perhaps this point of view belongs to Freud, if you want to attribute it to anybody.

The second view might be that personality develops of its own accord; then some-

where along the line, at a critical juncture, a blight occurs and impedes the growth of the personality. Something happens which makes the personality deviate. This might be labeled as Adolf Meyer's point of view.

The third might be that personality and psychopathology are different and never the twain shall meet. They are independent of each other. This might possibly represent Kraepelin's point of view. If at the present time you look over the literature to see which of these models is most acceptable, there is nothing to choose between. There is as much evidence for the first, second, and third; it is indeterminate. Perhaps, the third point of view is most acceptable to psychologists since it represents the null hypothesis, so frequently adopted in testing hypotheses.

If there is no relationship between personality and psychosis, why does Dr. Katz stress personality in this picture? This comes from some published work of which he and I are coauthors, and I wanted to remind him of it. (Joseph Zubin and Martin Katz, "Psychopharmacology and Personality." In D. Byrne and P. Worchel (editors), *Personality Change*. New York: John Wiley & Sons, 1964.)

Let's consider for a moment the situation where you have a drug working on an individual. On the one hand you have the drug dosage. On the other hand you have performance in a given function—let's say it is a reaction time experiment. When there is no dosage at all, his behavior will be determined by his usual way of behaving, namely, his personality.

I define personality as the characteristic way a person behaves when there is no interference or anything else impinging upon him. As the dosage increases you get a dosage curve relating performance to dosage and finally you reach a point where the drug takes over and everybody will behave the same way. (They will all fall asleep, or they will all become hallucinative, for example.) Somewhere in between there is a flexion point. Maybe this point might

be an interesting point to look at as a basis for classifying different personalities with reference to typology; maybe some day we can get a pattern here which will be indicative of the personality.

The same thing might be true of disease. If one ever develops schizophrenia, it might take a very short time before his personality is overcome at this inflection point and the disease takes over. With somebody else, it might take longer. You can apply the same principle. To a certain extent the illness will depend in its expression, in its timing, on the kind of personality one has and this may also predict what will happen if he becomes sick. It may be prognostic. So even though you might say personality and psychopathology are independent, it looks as though when you consider the results of an illness it is important to know the premorbid personality, since it may predict outcome.

Dr. GRINKER. I shall comment first, and then my coauthor, Dr. Nunnally, will pick up in his special area of competence.

I spent so much time talking about observations simply because these are the source of data. Unless the observations are good, the data and all of its analysis will be no good. As of now we have found the best observers of behavior are the subprofessionals and the nurses within a psychiatric unit.

The question of linearity which has been raised is a puzzling one to me. I doubt very much, at least in the field of mentation, whether there are any phenomena that are linear in nature. An example is anxiety. In lesser amounts anxiety is facilitative; in higher amounts it is destructive.

The criteria of typologies being valuable only for the analysis of drug effects is a misconception. Dr. Nunnally will talk about that later. The therapeutic aspect of drugs brings in further complications by all kinds of interferences which then make it very difficult to make the necessary correlations.

I was astonished that data in some papers show that each method apparently utilized very thin cross sections of behavior. To interview a patient once, to interview a relative once, makes it very difficult to elicit types having any meaning themselves.

Regarding Dr. Katz's paper: Should personality theory deal with dichotomies, to say there is a trend "in" or "out," (externalization and internalization), to follow simpleminded notions of a dichotomy? It would seem to add very little to an understanding of personality. There are many other ways of understanding personality, and this is what we must do in establishing our typologies. Diagnostic labels are of very little help. The notion of disease entity is of very little help. Certainly we must develop a better picture of personality. Basic data of behavior are most important for the development of meaningful types.

Dr. NUNNALLY. In my joint paper with Dr. Grinker considerable time is spent on the purpose of employing nosological classifications. We are depressed with respect to certain types of mathematical models implicit within much of the psychiatric diagnosis and classification today. Diagnosis today is used mainly as an inefficient halfway house between very good data and the prediction of important outcomes about patients—outcomes in terms of the mental changes over time; outcomes as functions of different types of chemotherapy, psychotherapy, and so on.

Logically diagnosis needs to be done for many different purposes, not just one. Unfortunately, we seldom stop to ask ourselves what is the purpose of diagnosing at all. If the diagnosis is for the prediction of some sort of plot outcomes, then it is far better not to reduce your data to types; in so doing you surely are going to lose a great deal of information.

There are instances in which one might want to reduce the data to types. One example is when you want to reduce the total collection of data to a small number of

types relating to your psychological and psychiatric theories, not to particular plot outcome. In this case it may be very useful to throw away information you simply cannot handle in your theoretical system.

Also, I am equally in favor of testing hypotheses, clinical hypotheses, about the existence of different types.

Let me make several comments on points made by Dr. Greenhouse. In this instance I am going to have to rebut my coauthor on the importance of nonlinear relations. Nonlinear relations in psychology are so few it is extremely difficult to find a good example of one for a textbook. When you find such nonlinear relationships they are seldom between psychological traits such as anxiety, introversion, intelligence, and so on; rather, they are between individual psychological traits or factors among such psychological traits and nonpsychological traits, certain kinds of performance data in learning situations, for example. You very seldom find marked curvilinear relationships between measures of psychological traits.

Even if there were such nonlinear relationships as a theoretical danger, the correlation between psychological traits is so low you couldn't possibly obtain much of a curve in the relationship.

I will agree with a number of you who said that until we get better data we probably should not apply more refined methods than we have.

The methods we already have, assuming the linear relationships, take too much advantage of chance. If we allowed more parameters to enter these equations in factor analysis and so on, we could prove almost anything we wanted to about proper manipulations of data.

Dr. LORR. Dr. Lehmann's paper implied that our research reflects on psychiatric competence. We are only trying to set up a research scheme which, hopefully, will be useful. He also asserts that the psychiatric diagnostic system is well established. Well,

is it really? I suggest that he read Dr. Menninger's book, if he is so convinced of that.

I would argue what we have done is something very much akin to what Dr. Grinker did. We collected data on rather sizeable groups of patients diagnosed as functional psychotics by diplomates in clinical psychology and psychiatry all over the country. We were very careful that the patients were not placed on drugs. We ran careful hospital-by-hospital checks to assure us that the raters were in reasonable agreement. Actually the agreement was extremely high. I know that in cooperative studies you often do get shoddy data. We tried to minimize this.

In developing our scales we did cover the literature very carefully: All factorial studies, the APA nomenclature, and current textbooks.

We have not claimed that our categories represent a set of psychiatric diagnoses. We are first trying to get a cross-sectional picture of the patient. We then plan to relate our categories to social history and to outcomes. And at that stage maybe we will be able to approach something like a diagnosis. But surely we are not suggesting that our types supplant current psychiatric categories.

Dr. OVERALL. I would like to say a word in support of what is being done, and can be done, with factor analysis, cluster analysis, and related procedures since these techniques seem to have come under rather severe criticism in this meeting. For very pragmatic and practical reasons it is important at this time to have a most parsimonious system of classification. For example, in the clinical drug research being supported both by NIH and other sources most sample sizes are very small. If we are going to make any headway whatsoever in identifying differential indications for psychotropic drugs, it is necessary to have a few very basic, perhaps initially gross, classification categories to work with. With small patient samples, fragmentation into more than two or three subgroups leaves

you with hopelessly small numbers. To have too broad a range of classification categories would be so confusing in terms of evaluating differential drug response that it would be a hopeless situation until we at least have the basic distinctions pinned down.

Factor analysis is a very powerful method of reducing dimensionality of the total symptom or total measurement space into its most parsimonious and meaningful form.

The cluster-analysis techniques, in addition, have the advantage of identifying areas of high density, or high concentration of patients, within this reduced measurement space. It is important in defining classification regions within the space that the classification boundaries do not cut across modal clusters, or the areas of high density in the space, if we are to achieve high reliability in our classification of patients. Obviously, fewer errors will be made if most patients tend to fall well within the center of classification regions and do not tend to straddle the boundary lines.

Another advantage is that the cluster types tend to be defined in terms of objectively measurable characteristics and hence are amenable to objective rules for assignment of patients.

Major clinical types appear entirely inadequate for identifying differential responsiveness to treatment, and have not proven useful in the past for treatment purposes, prognosis, and so on. And I think that is one reason we are here at this conference today.

The cluster procedures, the factor-analysis procedures, and the other methods aimed at identifying relatively small numbers of major classification regions with the measurement space, independent of specific prediction of drug response, are very important. We have a large number of drugs at the present time, and we have good reason to suspect there will be more to come along in the future. We would soon end up in a

hopeless confusion with specific functions aimed at predicting response to each specific drug.

Partitioning of the measurement space in a meaningful way—which, at the same time, has usefulness for purposes of evaluating treatment—is a very good recommendation at this time.

These are my comments as to why I think we should continue to use the factor-analysis and cluster methodologies that we have been using.

Dr. Lehmann complains of the time required to train residents in the evaluation of even the most parsimonious set of perhaps abstract symptom and behavior concepts, such as guilt feelings and anxiety. I would ask, in response to this: How many hours of training are consumed by descriptions of traditional clinical diagnostic types, in the hope of bringing the student to the point of being able to diagnose mental illness using these more complex concepts? Certainly it would take less time to clarify standard symptom-descriptive concepts and to train residents in reliable means of evaluating a relatively small number of distinct symptom and behavior characteristics.

Dr. Lubin, in discussing our work with regard to its application to differences in drug response, made the statement in his paper that patients were randomly assigned to the drug treatments. I will read the paragraph we have in our paper relative to that.

It says:

While the several drug treatment groups were not studied concurrently, the studies followed a similar general protocol with sampling from the same patient population. Patients in the studies of antipsychotic drugs can be considered to represent random samples from the population of newly admitted (or readmitted) male schizophrenic patients in the Veterans Administration. Antidepressant drugs have been studied in clinically depressed patients. The research design has routinely called for pre- and post-treatment ratings by two independent observers

using the 16-variable brief psychiatric rating scale.

A very cogent criticism, and one that this paragraph was inserted to allow the reader to evaluate, is that these samples were not randomly drawn from this population at the same time. They were drawn according to the same instructions, and according to the same selection criteria; they were drawn from the same hospitals, and evaluated in general by the same raters, although there were very few changes in raters over the period of time these studies were collected.

So the time dimension does differ and we can only assume that the population from which these samples were drawn remained essentially the same during the period of these investigations. I wanted to comment on that point and recognize that problem. And I thought it had been explicitly recognized here.

Dr. KATZ. Our work, the data we collect is not, contrary to Dr. Grinker's point, a thin slice of the cross-sectional behavior of patients. Informants who have been in a position to observe are asked to describe the nature of the patient's behavior and symptoms during the past month. It is because an informant reports on observed behavior over a period of time that a type like the "periodically agitated withdrawn" can be identified. The designation refers to the fact that this type is basically withdrawn but has periods of extreme restlessness or agitation.

In assessing the value of a particular typology, Dr. Lubin places great emphasis on the importance of predictive validity, the typology's potential for predicting treatment results, and considers its consensual validity as irrelevant. We are very far apart on that. I not only consider the latter relevant, but I view it as the primary concern where validity is at issue. Whether a typology can predict in any specific instance, e.g., with regard to a particular treatment, has to be a secondary concern, and should theoretically depend on whether the typology exists in the first place. I ex-

plained in detail in the paper why that was true—but it's worth repeating that the long-term theoretical and practical value of any new typology depends on how real it is—how likely it is that it reflects the true state of affairs—i.e., the several configural patterns in which psychopathology is actually manifested.

Dr. Greenhouse defended the dropping of cases with regard to the chlorpromazine analysis. This had to do with testing whether two different patient types responded differently to two different drugs. His defense of that was not complete in the way I saw it; I would like to defend it myself in a different way.

The hypothesis to be tested was whether two patient types would respond differently to two drugs. I think that when one deals with the testing of a particular hypothesis, his only concerns should be that the data and the measurement be appropriate to the testing of that hypothesis. The fact that other data is also available to answer other questions, and that this data has to be ignored or thrown away, as Dr. Lubin puts it, may disturb one's sense of economics, but it obviously shouldn't be permitted to confuse the research aims or influence the design of the experiment.

With regard to why we dropped chlorpromazine, that was indicated in the paper. We simply didn't have enough cases to include it in any types by treatment analysis.

Dr. Lubin also chides me about shifting from one measure of improvement to the other. I plead guilty there. There are several facets to the concept of improvement, and many ways of measuring it. There has been much study, much controversy, and many theories as to what improvement is. We had a problem when we moved to the types by treatment situation where we were forced to work with very small groups of patients. Because of the small sample size, there is then a need for greater precision, and improvement measures have to be se-

lected which are more reliable than those which can be used in studies where very large numbers of subjects are available and where more error can be tolerated. In this case, in selecting the second order factor of schizophrenic disorganization from the IMPS, we moved to as reliable and as relevant a measure as we could find with regard to this particular problem.

The problem of which improvement or change measures are appropriate has to do with the kinds of experimental problems that arise when you wish to test this kind of interactive hypothesis; that is, where there are two different types of patients who respond equivalently, say, to a drug. You are starting out with the assumption, and have demonstrated, that these are two very different types of patients. Is it reasonable then to expect that they will improve in the same way or change on the same behavioral parameters?

In seeking a common baseline, so that it can be determined whether two very different types are equally changed or improved, then it turns out as you'd expect, that it is very difficult to find behavioral or symptomatic measures on which they are alike to start with. If you could find such measures, you probably would be wrong about the typology in the first place.

I don't say that in the study I reported that the problem was solved. It is still a very difficult problem as far as our own research is concerned, and one for which no simple solution has been worked out.

I want to say to Dr. Lehmann: Research problems in the clinical field require more objective methods than we have been using in the past, and more objective ways of classifying patients, and the techniques that have been described are designed to facilitate certain kinds of research in the field—and not to substitute for clinicians or clinical practices. Unfortunately, there is wide variation among clinicians with regard to the quality of their observations and the quality of their judgment. In large studies where large groups of clinicians with varied

amounts of training and experience are involved, we have to apply more systematic, hopefully more objective methods in order to have a reasonable chance of meeting the aims of the research.

With regard to his comments on our results, he said he could have predicted that these particular kinds of patients would have responded to these kinds of drugs. I was surprised by the results. I had asked several clinicians, before the fact, what they would have predicted with regard to these drugs. I asked, for example, about how this particular drug would affect a belligerent manic patient, and the general consensus was that it would do fairly well with that type of patient; and that it would also do well with the acute panic type patient. Well, the result as reported is that the acute panic patient does do well, but the belligerent manic does not. I don't think there is as much agreement as Dr. Lehmann indicated there was with regard to whom these drugs would help—and, thus, more evidence that objective approaches have something to tell us about how to solve these problems.

Dr. BALL. All of these studies went through the data once. I wonder if anyone in these studies has gone back to the patients, perhaps breaking the groups up even more finely, interviewing the patients within each of these smaller groups sequentially. Since the clinician would be observing each relatively homogeneous subgroup of people one after another over a shorttime interval, it might be hoped that previously unnoticed common characteristics of these subgroups would become obvious. This process could be repeated several times with each cycle hopefully approaching more closely a useful set of measurements.

Dr. LORR. I think Nils Mattson has analyzed repeated measures on the same patients. I understand Cattell is doing this in the case of depressives. I have always felt that when you define a class of patients, if you then go beyond this and apply other devices to

the patient, you get a broader spectrum of behavior. For example, by getting more verbal self-reports from anxious depressives it is possible to secure greater differentiation among them.

Dr. HAMILTON. This morning we dealt with the theoretical basis of classifications, and this afternoon we have been concerned with actual investigations which, presumably, were brought here to exemplify these particular theoretical approaches. It is for this reason that I intend to discuss the specific details of two of the papers which have been presented here; or, at any rate, the printed version of them. Much has been made this afternoon about the question of good data. I am not quite certain why. Good data are like motherhood: We are all in favor of it. Good data are described as being those which are reliable and valid; but they also have other properties such as meaning and relevance.

Now let me take an awful example first. In the paper by Dr. Overall he started off by asking 38 experts to define the various groups. This investigation is not on schizophrenia or mental disorders; it is on what the experts think about them. It may be very interesting to investigate the psychology of experts, but this is not the same as investigating mental disorder. Nevertheless, although at first he started off with the stereotypes of experts and the APA classification, and then went on to deal with empirically derived data, he very quickly discovered the dichotomy of depressions; i.e., the retarded versus the anxious, agitated types. These types of depression, which have been known and argued about by psychiatrists for something like 50 years, were not included in the original stereotypes.

Drs. Grinker and Nunnally started off by deploring the preoccupation with dynamics and the absence of sound clinical observations; but they are still bound by this preoccupation, for the data that they have used for their work consists of feelings and concerns, plus current behavior. Why did

they not deal with symptoms? This is not an irrelevant question; among the factors that they found, and the defining variables of these factors, I notice the absence of one of the most important symptoms of depression: suicidal tendencies. Although it is not mentioned in the paper, in their original work there is only 1 item out of well over 40 which asked this question. I would have thought it was worth a few more.

I noticed that among their factors there were not included loss of weight, loss of libido, and so on. These are very important symptoms, not only for the diagnosis, but also the treatment of depressive disorders. The net result of the analysis of their many items or questions is that they came out with 15 factors. It is not an accident that this number is of the same order as the symptoms of depressive states. In my work I used 21, of which I analyzed only 17 for various technical reasons.

Contrariwise, when they worked with items concerning clinical behavior they come out with 10 factors; and I can only say I cannot tell the difference between one and the other. They are so overlapping that it seems to me that they are describing very little more than four particular types of behavior, and these are our old friends the symptoms once again.

It is important that we should use every method possible for examining and observing patients, and analyzing the data; but it is fundamental that we should work on the basis of experience already obtained. This experience shows its importance by the fact that if one ignores it, then after much work and effort, one comes back to it again.

One final point I would like to make about this paper: there is much emphasis on what they call saliency. I am not quite sure that I understand what precisely is meant; whether it is meant to be a dominant symptom or a leading symptom, or what I would call a pathognomonic symptom. What is clear from the paper is that there exists a somewhat simple-minded approach (not made by the authors, but by

others) that if a patient is very depressed he will be diagnosed as suffering from a depressive state; whereas, if he has, say, hallucinations and delusions, he will be diagnosed as schizophrenic. But surely nobody makes a diagnosis on one symptom! One doesn't even make a diagnosis on many symptoms. A diagnosis involves much more than merely 1, 2 or even 10 symptoms. Any one of them may be absent, and yet the diagnosis will be the same. It will be based on the course, the development, the background, and the final outcome. All of these are involved; if we are going to use diagnosis merely as a label based upon one symptom, we will end nowhere.

Dr. NUNNALLY. I appreciate Dr. Hamilton's concern with symptoms. We surveyed the entire literature on depression and ended up with over 300 specific symptoms, all the way from dry mouth to constipation and loss of weight. But in our early investigations we found most of these very seldom occurred, or at least were observed to occur, in depressed patients.

I am glad you brought up the point about suicide; we were very surprised, too, because this was not observed either in terms of any overt thing the individual had attempted or in terms of a diagnosis on the part of the psychiatrist. So what we ended up with were about 15 or 20 symptoms, as one would usually speak of them, in our current behavior sample.

So there were symptoms, and they did crop up. But the surprising thing to us was that the classic symptoms one reads about in the literature apparently either very seldom occur in people spoken of as being depressed, or they really are not accurately observed by the psychiatrist in the situation.

Dr. OVERALL. I am glad Dr. Hamilton pointed out that we had to discover retarded depression for ourselves. The prototypes we used in the initial investigation were based upon the major categories from the standard nomenclature. I believe that the

distinction between phenomenologically different groups of retarded and agitated depression does not appear in the standard nomenclature.

Dr. KLERMAN. I should like to respond to the comment on Dr. Greenhouse. Usually, as one of the younger clinicians, I find myself arguing with my clinical elders. Today, I shall reverse the approach, much to my surprise, and agree heartily with my elders.

I think our problem is a failure to communicate an understanding of the traditional medical concept of diagnosis of disease. I do not know whether this is due to our colleagues in psychological statistics who have not read our history, or whether it is caused by our failure to make it clear.

The concept of disease does not mean a covariance of symptoms occurring at one point in time, or even over a brief period such as a month. The history of this work today, it seems to me, is a recapitulation of my understanding of the early 19th century. In the first half of the 19th century, there were many systems based on clustering of symptoms. There was a plethora of them. They were overlapping, almost to the same extent as the excellent typologies presented by the investigators today overlap. The problem that the later 19th-century—mostly German and French—investigators faced could be expressed in a question: Was there a way to make sense of the plethora of symptom complexes at one point in time?

The principle that Kraepelin propounded was to apply to psychiatric symptoms the principle evolved from general pathology, as exemplified in the work of Virchow. The attempt was to classify, not on the basis of any symptomatic clustering at one point in time, but on the basis of etiology. Since then, the basic principle of medical classification is an etiological principle. This was a new historical view, as historians of medicine such as Rosen and Shyrock have pointed out. The concept of disease classification based on etiology is a

product of the late 18th century and was only validated by the work of bacteriologists and the pathologists in the mid-19th century.

Kraepelin not only contributed the delineation of kinds of dementia praecox, but he also was one of the first to apply to psychiatry a pathological and etiological classification. Those who read his general textbooks are impressed, as I have been, that he had delimited categories of psychosis due to head injuries, alcohol, and drugs. And he thus applied to the psychiatric phenomena the psychiatric symptom clusterings postulated by a previous generation. He was then faced with the problem that there remained a residual category of patients with chronic psychoses for which there was no apparent biological corollary. He then shifted gears and attempted to say, "Can we make sense of this group for which there is no apparent biological correlate by some other criteria?" He applied the criteria of course over time and separated the residual category into two major subgroups: (1) The manic depressive, and (2) the dementia praecox.

It seems to me that a great part of our communication difficulty results from the failure of clinicians to make explicit our historical traditions. There is a danger that the work on statistical nosology will recapitulate the mid-19th-century confusion.

Is the principle of biological causation irrelevant to psychiatry? Many people argue that it is. Dr. Shakow made this point yesterday, and Dr. Menninger has argued it very eloquently in his recent book, "The Vital Balance." I do not think so. I think if one looks at contemporary research in mental retardation, for example, one finds brilliant progress where the principle of biological causation is bearing fruit. I wonder if we are in danger of giving up this principle too readily in psychiatry.

Even if one takes the point of view that there are relevant causal factors other than biological ones, such as familial background or early childhood development, one still cannot, I think, avoid the concept that disease is more than just a covariance of symptoms at one time. Those who believe, as most of our colleagues do, that the cause of schizophrenia is early childhood development, or even symptomatic of some familial pathology, assume the same Kraepelian concept; they have shifted the locus of pathology out of the head into the family.

Perhaps, in addition to sniping at each other, clinicians need to do more about making explicit to our colleagues some of our traditional concepts.

Biometric Assessment of Mental Patients¹

Joseph Zubin, Ph.D.²

Perhaps it is necessary at the very start to draw attention to the nature of the biometric approach before launching into the problem of assessment. The biometric approach is to be distinguished, on the one hand, from the taxonomic, and on the other, from the statistical approach although it shares common ground with both of these. Biometrics is the science which applies measurement to living organisms—including, of course, man. Taxonomy is concerned primarily with classification of organisms, objects, or entities into categories that are useful and meaningful such as species, genera, diseases, etc. Statistics, of course, is a tool for ordering data and drawing inferences from the obtained orders. Biometrics provides measurements which can be used by taxonomists and which can be analyzed and synthesized by statisticians, but the substantive contribution of biometrics ranges beyond these two applications. It forms an objective basis for aetiological investigations, prognosis, and discovering of new taxonomic categories, as well as the investigation of the relationships within and between organisms in the obtained measures.

While taxonomists utilize biometric data,

the nature of such measurement is usually on a primitive level (either nominal or at most ordinal scaling). The distinction between biometrics and taxonomy can be put in the following terms: The biometrician collects data in order to establish diagnoses while the taxonomist collects diagnoses as data by which to justify the categories he has chosen. Thus the biometrician provides data for assessing patients while the taxonomist provides data for assessing diagnosticians. Both methods are valuable, but the distinction must be drawn in order to understand why the techniques and methods may differ in accordance with the goal. Thus while the taxonomist or diagnostician, biostatistician, and biometrician each has a place at this table, their functions and purposes are not identical by any means and bearing these differences in mind, it will be easier to understand the contributions each will make.

Perhaps the following contrast provided by Karl Jaspers will make the distinction even more trenchant. In this passage he refers to the clinical evaluation of body types à la Kretschmer, but it is equally applicable to other clinical evaluations. (Jaspers, 1963)

"The biometric methods give us more than figures and correlations. They foster clarity in all fields in which biometric variations can be established. Moreover, through the application of these methods we have concrete experiences which we would never have had without them

¹ The preparation of this paper was aided in part by Grants MH09191-02, MH07477-03, MH03546-05, MH07776-02, MH08534-02, MH07478-03 of the NIMH and GS-611 from the NSF. The author wishes to thank his colleagues in Biometrics Research—E. I. Burdock, Joseph Fleiss, Kurt Salzinger, Robert Spitzer, and Samuel Sutton for their permission to quote the results of research under their guidance and for assisting in the writing of this paper. The author is indebted to Mr. David Jenness for editorial assistance.

² Chief of Psychiatric Research (Biometrics), New York State Department of Mental Hygiene, New York, N.Y., and Professor of Psychology, Columbia University.

though they may once more be lost to sight in the purely statistical results.

“* * * As we proceed we acquire concepts and ways of looking at things which in fact are unanswerable questions. They are the key to wide possibilities but they bring us no definite knowledge. Our mind is kept open for the true whole in its present concrete form as an experience which constantly gains in depth. It is as if we could tangibly grasp it but it evades our clutch while it never opposes our efforts to penetrate yet further and advance.”

Thus, while accentuating the need for an evaluation of the totality of the patient, Jaspers nevertheless points out the need for specific measurement as a means of clarifying the total picture.

Why the opposition to measurement persisted and still persists today is one of the paradoxes of science. Perhaps one of the reasons is the tradition, firmly entrenched, that only physicians can adequately study medical problems. This is tantamount to saying that ichthyology should be turned over to fishermen, sociology must be studied only by social workers (Shryock) or astronomy by sea captains (Comte). Fortunately, the present conference is not hampered by such considerations.

Is the biometric method new in psychopathology? Far from it. Pinel was one of the first to use statistical methods (1809–20), and Kraepelin in 1890 had this to say about biometrics itself:

“As soon as our methodology has sufficiently proved itself through experience with healthy individuals, it would be possible to approach the actual ultimate goal of these efforts, the investigation of sick personality, especially of the inborn pathological disposition. In an investigation of many individuals we will always find some who deviate profoundly from the behavior of the vast majority in one or another aspect. If this deviation appears to be damaging to the mental life, and if it reaches a certain degree—which admittedly

can only be arbitrarily determined—then we tend to regard it as an illness. Experience teaches us that persons with pathological traits of this kind are, on the whole, in greater danger of a general mental disturbance than those personalities (natures) whose characteristics are in the middle range. We therefore have first of all to investigate whether it is possible by means of psychological tests to determine individual deviations which cannot be recognized by ordinary observations. If that succeeds, we would be in the position, through the quantitative determinations at our disposal, to establish the borderline between health and disease much more precisely and more validly than has been possible so far (1896, p. 77).”

Having differentiated between the clinical and the biometric approach, let us now examine the present status of diagnoses from the point of view of both methods.

There is an uneasy feeling abroad that there is something wrong with our diagnostic system. Whether this is indeed the case or not is difficult to determine but that there has been a loss of interest in the making of diagnoses and a consequent drop in their accuracy and value is undeniably true. The cause of this loss of respect for the diagnostic system has had many sources. Among them are, first of all, the long period during which diagnoses were purely of academic interest because only one treatment—namely, custodial care—was available; second, the repeated demonstration of the lack of reliability of the diagnoses; and third, the low relationships that were discovered with regard to its validity as measured by the course of illness, outcome, and relation to treatment. It is disheartening to note that as the armamentarium of new therapies expanded, and the need for better classification and diagnosis increased as a consequence, the acumen and ability of the diagnosticians seemed to drop.

To counteract this general trend which has led in the extreme to the denial of the

need for diagnosis and to the claim that mental illness is a myth, a new approach to the evaluation of the patient's behavior is essential. While the current methods and techniques in the hands of gifted clinicians yield insight into the behavior of the patient sufficient for clinical treatment, they are unsatisfactory for research purposes. The biometric approach takes as its point of departure the utilization of objective methods in the assessment of patient behavior without regard to current diagnoses. The chief purpose is to document the overt behavior of the patient as well as his mental content in order to provide objective data on the basis of which classification and effectiveness of treatment can be made.

Classification is an essential step in the evaluation of treatment, since evaluation can be made on a group basis only, and the provision of homogeneous groups is the *sine qua non* of good classification. To this end the entire spectrum of behavior is examined and use is made of the entire range of disciplines from the physiological to the anthropological in obtaining information about the characteristics of the patient. By means of such a wide-angled lens the entire behavior of patients can be viewed and methods for classifying individuals can be obtained on the basis of a variety of behaviors, each of which has some implication for assessment.

The attempt at sidestepping diagnosis and developing techniques for the objective assessment of behavior without reference to current diagnosis has certain dangers. It tends to ignore the current grouping of patients and the classifications which lead to the well known diagnostic categories in the mental disorders. However it also has certain values. It permits the observer to view the behavior without bias and permits the data to be grouped independently. If the current classification categories are educed, well and good. If new classification categories emerge, it would be a gain for science.

The biometric assessment of mental patients involves objective techniques and methods that are less dependent on the subjective and intuitive skills of the clinicians than are the usual methods of clinical evaluation. Where subjective methods have to be resorted to, an attempt is made to make them explicit and public so that others can replicate the obtained data wherever possible. Where no such replication is possible, appeal is made to a consensus on the part of trained observers. Thus, even if we return to the current nosology, we will do so on the basis of biometric data and not on the basis of the subjective, intuitive, albeit insightful clinical approach.

However, before assessment can begin, we must bear in mind some type of framework against which the assessment is to be made. The number of characteristics, traits, and behaviors which man is capable of is infinite and the assessment must selectively be based on some theoretical framework or scientific model. The scientific models which have proved useful in the field of psychopathology are the following: (1) The social-cultural model, (2) the developmental model, (3) the conditioning or learning model, (4) the genetic model, (5) the internal environment model, and (6) the neurophysiological or brain function model. Until recently, the first three models—social-cultural, developmental, and learning, were the most prominent. In recent years, genetics, internal-environment, and brain-function models have become more popular.

Before introducing the aetiological models it might be well to point out that at the present time the most generally accepted approach to psychopathology is the naturalistic approach of description of behavior. This is always the forerunner of more specific models, since without a knowledge of the phenomena no aetiology can be understood. We have had several hundred years of description by astute observers and much knowledge has been

gathered. On the basis of these observations, a rather loosely organized structure has been developed, in which observation and aetiological inferences have been intertwined. It is time now to undertake a critical evaluation of the tenability of the structure and of the validity of the inferences.

The social-cultural model is built on the assumption that all mankind is vulnerable to mental disorders and that given sufficient deprivation, stress-producing loads, or other alterations in our environment, our behavior will be altered to the point where our ability to continue living normally as independent individuals in society is endangered. The evidence for social-cultural-environmental pressures as aetiological agents comes largely from studies of socioeconomic status, isolation, educational and social deprivation and social-cultural uprooting in immigration or migration or rapid acculturation. Even the most sanguine environmentalist will not be satisfied with merely pointing to the above mentioned factors as "causal" agents, but will try to determine just how these malignant factors bring about their deleterious effect. While the story is far from told, there is already sufficient evidence to at least question whether these factors cause mental disorders.

A more thorough review of the evidence which has been presented elsewhere (Zubin, 1963) leads to the belief that social-cultural forces may elicit a mental disorder or even occlude it, but cannot cause it unaided. However, much more research is required to transform this belief into fact.

Despite this conclusion, the fact remains that the detection, diagnosis, and even rehabilitation of the mentally ill today rests on a social-cultural framework, deviations from which identify the mentally ill, and return to which constitutes the basis for improvement.

In summarizing the social-cultural model, we might point out that the cur-

rent revolution in management in psychopathology with regard to hopefulness of treatment, reduction of patient population, rehabilitation, etc., is to a considerable extent a social-cultural change involving change of attitude on the part of patient, family and therapist. Hence, though social-cultural forces may not be so important in aetiology of some disorders they are of great importance in detection, treatment, and rehabilitation.

The developmental model for aetiology is built on the assumption that mental disease develops as a result of some specific deprivation or interference during a critical period in development when the specific deficit or interference is crucial. Evidence for this model is afforded by the recent investigation of Pasamanick (1961) and his collaborators on the role of intrauterine events on the continuum of "reproductive casualty." They postulate that certain untoward events, such as intercurrent illness, toxemia, and other interference with the foetus during the first 9 months of life will produce mental and physical disability ranging from stillbirth, through live births with epilepsy, cerebral palsy, mental deficiency. Finally, even those who appear unscathed at first may not escape entirely but develop such lesser ailments as reading disability. One of the most exciting events in the developmental area has been the investigation of the impact of early experience on subsequent personality and its deviations. While the evidence from human infants remained controversial, the evidence from animal studies seemed at first to yield data which appeared too closely in keeping with Freud's hunches to give much comfort to those who had refused to accept his clinical surmises. Gentling of animals in their infancy, following the credo of tender love and care, actually produced less emotionally unstable adults, but to the great surprise of most investigators, shocking the infants was equally effective and most recently, Theodore Schaefer et al. at Columbia (1962) has found that merely lowering

the temperature a few degrees is equally effective. Thompson's (1962) demonstration that emotional mother rats give rise to emotional pups, Harlow's (1962) demonstration that monkeys raised on surrogate mothers and not permitted to play with their peers tend to develop poorly in the psychosexual sphere, and Melzack's deprived dogs (1957) are still other triumphs for the developmental model as a possible causal factor in emotional disorders.

The conditioning or learning model postulates that the source of the deviant behavior of the mental patient is to be sought in his reinforcement history. An example of this theory is Bateson's (1956) double-bind model in which the mother's ambivalence in her relationship to her offspring evokes ambivalent behavior and other types of deviation in him which we recognize as schizophrenia. While Bateson's double-bind model has aroused considerable interest in psychodynamic circles, it has thus far defied experimental testing of any of its hypotheses. Several more experimentally founded models have been provided by psychologists. Thus, Sarnoff Mednick (1958) bases his approach on the evidence that the early or acute schizophrenic conditions more quickly, and shows greater stimulus generalization (less steep gradients). These are related to the higher level of arousal which is attributed to early schizophrenia.

With regard to the genetic model, it should be recognized that strictly speaking there are probably no exclusively genetic or environmental disorders. All disorders are both genetically based and environmentally elicited. Without the hereditary-environmental interaction, no disease, in fact no development at all, would be possible. What then is meant by a hereditary disease and by an environmental disease? PKU is a hereditary disease in our particular social-cultural-physical environment because of the presence of phenylalanine in our diet. Had our diet been free of this substance the phenylketonurics in our so-

ciety would never develop mental deficiency and, in fact, PKU would never have been discovered. On the other hand, if only the poor, or only the mountaineers had developed this illness, and the error of metabolism were unknown, we would have regarded this illness as environmentally produced.

Thus, a disease is regarded as definitely hereditary if we already have evidence of the presence of the specific genetic component and have not yet discovered the specific pattern of environmental components required for eliciting it. On the other hand, a disease is regarded as environmental if we have found the specific pattern of environmental components but have not yet discovered the pertinent genetic factors. In diseases where both the hereditary and environmental aetiology is known the question never arises. In the end, however, all diseases will probably be found to require both a genetic as well as environmental component. Apparently, heredity is no more the cause of an illness than the automobile is the cause of an automobile accident. The interaction between specific hereditary and environmental factors required for the emergence of an illness must be sought if we are to detect the vulnerability of an individual before the illness overcomes him.

Genetics may be viewed as a biochemical mechanism in which the genes serve as precursors for the production of certain enzymes whose absence prevents the organism from prospering. There is, therefore, considerable hope that an investigation of the internal environment of the body may reveal the particular metabolic deficiency or excess which characterizes the patient. A particular error of metabolism may, of course, be inherited or acquired. A considerable amount of effort has been spent in the attempt to relate schizophrenia to metabolic error. Certain fractions of schizophrenic blood have produced metabolic changes and changes in such behavior

as rope climbing in rats as well as transitory changes in the psychomotor behavior of normal human subjects. Presumably similar fractions from the blood of normals do not produce such changes.

Perhaps the most exciting new development is the discovery of the ability to manipulate behavior directly through implanted electrodes. This has given new significance to the work of the neurosurgeons who carried out psychosurgery in the late forties and early fifties. The attempt to detect deviations in the neurophysiological substrate of mental patients through the means of electrodes implanted in various portions of the brain, by recording evoked potential or by introducing current through these electrodes, and the recording of evoked potentials from the intact brain through the scalp, are opening up new possibilities for the detection of the neurophysiological anomalies correlated with deviant behavior.

The final scientific model is the epidemiological model, which seems to be a super-model including each of the above models as partial factors in the explanation of the mental disorders but requiring careful field studies to determine the relative role of each of them. Mental disorder is conceived as the end result of a series of probabilistic events, each of which must occur in interaction with others to produce the disorder, although the threshold value for each factor may differ from person to person and from one disorder to another. Thus, two people may have inherited the same predisposition, but because of differential stress, nutritional or deprivational factors, will not both develop the illness. The virtue of epidemiology is that it takes in all possible factors ranging from radiation, paranatal existence, genetics, to social-cultural environment, etc. Thus, the epidemiological model both permits and requires the weighing of each of the submodels in the total picture of causation; the difficulties of assessing their relative importance, and of devising studies which will not overlook

some of the factors, are too well known to need reemphasizing here.

While many of the social-cultural, developmental, learning, genetic, biochemical and neurophysiological claims remain somewhat in doubt, they nevertheless lend credence to the possibility that the mental disorders will be found to be characterized by either the deviations which are now postulated by these models, or by deviations of the same general scope that have not yet been postulated. Whether they are the cause or the effect of the disorder remains to be seen, but the testing of the hypotheses generated by these models depends to a large extent on the detection of some type of deviant behavior which characterizes the patient.

One question that is often raised regarding the behavior of the mentally deviant individual is the question of its continuity vs. discontinuity with normal behavior. This is an age-old issue which has never been dealt with conclusively. But it is interesting to inquire where the discontinuity, if it exists, is to be sought: in the patient or in the observer—i.e., in the behavior of the patient or in the behavior of the community with regard to the patient? If the discontinuity exists in the community so that, once a threshold of deviation is passed, the individual in question is regarded as abnormal, it becomes clear that we do not have to postulate a discontinuity in the patient's behavior. In fact the individual immediately adjacent to him on the continuum, who differs by only a small degree from him, may still be regarded as normal. Hence, it would be a vain hope to find universal characteristics of mental disorder, since they are detected only by deviation from local norms.

On the other hand, if discontinuities exist in the patient's behavior or in his characteristics (biochemical, metabolic, etc.) a search for such discontinuities would be rewarding. Another possibility to consider is that the discontinuity in the community's behavior or in the patient's behavior may

rest not on a single dimension but on patterns of several dimensions. In that case, no single dimension may exhibit any sharp discontinuity. Instead, the patterning of the dimensions may identify the disorder even when all the individual dimensions remain continuous.

Thus, an individual who has an IQ of 70 may still be within the normal range. One who has a social quotient of only 70 may also be within the normal range. But given the combination of IQ 70 and social quotient of 70, the individual may be regarded as retarded by a given community. Here too, it would be a vain hope ever to find universal patterns of mental disorder, since they can be detected only as deviations from local normative patterns. On the other hand, if discontinuities exist in the patient's behavior or in the patterning of his behavior, a search for such discontinuities ought to be rewarding.

Each of the six models described above demands different approaches to testing its hypotheses. Even the first model, that of naturalistic observation, involves some types of assumptions and special techniques.

The techniques which have been developed are of three types. First is the approach which deals with the culture dependent methods of detecting and diagnosing mental deviation. This can serve the naturalistic and the social cultural model. The second approach is the culture fair approach which can serve the developmental and learning and conditioning models. The third is the culture free approach which can serve the genetic, internal environment and neurophysiological models.

The techniques for measuring the culture dependent, culture fair and culture free techniques can be classified in a Mendelejeff-like table which relates stimulus to response as shown in table 1.

Table 1 demonstrates the stimulus and response classes of which functional analysis may be fruitful in studying psychopathology, under load and nonload conditions. The actual representatives of these classes—

the contents of the cells—have been and may be as numerous as the scope of research in psychology, psychiatry, and experimental psychopathology allows, and some areas of functional relations—"correlations of classes," as Skinner calls them—have not as yet been explored. Following this design, a classification of techniques for the study of schizophrenics has been offered by Burdock, Sutton, and Zubin (1958), and a similar sample of functional relations where psychopharmacological agents are used as "load" has been prepared by Zubin and Katz (1964).

Since these response categories do not exist in the pure state it becomes necessary to find some way of separating them experimentally or statistically. Some years ago we found that chronic schizophrenics whose conceptual test performance was higher than their perceptual test performance remained in the hospital on a 2-year followup while those whose conceptual test performance was below their perceptual test performance were discharged (Zubin et al., 1953). Our definition of conceptual tests and perceptual tests were common sense definitions based on the nature of the task in the test. Factor analysis (Pfeffer, unpublished; Machi, unpublished) supported these conclusions, but it became quite clear that some better definitions on a more operational level are required for separating conceptual from perceptual and sensory responses.

Recently we seem to have achieved a beginning in the direction of obtaining purer measures. The task put before the subject is the following: He is given a variety of pairs of stimuli (such as light or sound stimuli) in which the first member of the pair serves as a cueing stimulus which sets up certain expectancies about the second. Between the first and second stimuli the subject is asked to guess whether the second stimulus will be a light or a sound. The brain's electrical activity as recorded from the scalp is obtained. By suitable analysis of these evoked potentials it becomes pos-

TABLE 1.—Examples of Measurable Activities as Functions of Stimulus Variables

Level of observed behavior	Stimulus variables					
	Idling state		Energy variables			
	None		Appropriate energy		Inappropriate energy	
	Variable	Function	Variable	Function	Variable	Function
Conceptual.....		Reverie and fantasy.	Uniformly diffused light.	Fantasy.	Inversion of gravitational attraction.	Subjective account of experience.
Psychomotor.....		Spontaneous movement.	Painful stimulus.	Arm withdrawn.	Electroshock.	Movements of limbs.
Perceptual.....		Spatial and temporal orientation.	White noise.	Orientation to direction of sound.	Pressure stimu- lation above retina.	Phosphene.
Sensory.....		Background noises; cortical gray.	Light of graded intensity.	Threshold response.	Electrical stim- ulation of thermal receptors.	Warmth or cold sen- sation.
Physiological.....		BMR; basal EEG; basal PGR.	Increase in carbon dioxide con- centration.	Change in rate of respiration.	Pressure on carotid sinus.	Change in heart rate.

sible to detect regularities which can be shown to be related to the sensory component (modality), the conceptual component (the guess), and to the emotional component (effect of being right or wrong). By contrasting patients and normals in this experiment the relative deviation in the sensory, conceptual, and emotional components can be established (Sutton, Braren, and Zubin, 1965). While we have obtained data only on normals with this technique so far, the problem of extending it to patients is simply a matter of time.

But it is rarely possible to investigate all behavioral levels for all stimulus classes at once. While no single family of techniques is adequate to explore all areas of functioning in an individual or group of the mentally ill, we may profitably return to an examination of what we have called the culture-dependent, culture-fair, and culture-

free indicators to provide examples of the techniques that are being used to good effect.

Culture-Dependent Techniques

The culture-dependent techniques are based on the assumption that mental disorders reveal themselves through deviations from social-cultural norms. The initial detection of mental disorders is usually made by laymen—the patient himself, his family, friends, or public officials. Since the identifying behavior is usually unexpected or deviant, it is no wonder that different cultures will not recognize as mentally disordered the same kinds of behaviors. This may be the reason why such great differences are observed between cultures with regard to the incidence of various forms of disorder. Once the deviation is observed and the

TABLE 1.—Examples of Measurable Activities as Functions of Stimulus Variables—Continued

Level of observed behavior	Stimulus variables					
	Signal variables					
	Configurations		Signs		Symbols	
	Variable	Function	Variable	Function	Variable	Function
Conceptual.....	Aircraft forms or silhouettes.	Recognition of identity of forms.	Practical trouble shoot- ing test.	Diagnosis of trouble.	Stimulus word in associa- tion test.	Association to stimulus word.
Psychomotor....	Star-shaped maze.	Mirror tracing.	Classical de- layed re- sponse stimuli in animal ex- perimenta- tion.	Successful response by animal subject.	Reinforcement of affect in focused interview.	Electromyo- graphic response.
Perceptual.....	Rotating Benham Disc.	Subjective color ex- perience.	Usual visual alternatives in animal discrimi- nation exper- iment.	Selective response of animal subject.	Musical chords.	Pitch discrimi- nation.
Sensory.....	Patterned light stimuli.	Visual threshold.	Infant's faint cry.	Mother's auditory threshold.	Words or sentences presented tachisto- scopically.	Visual threshold.
Physiological....	Photic driving.	Change in EEG pat- tern.	Bell ringing in Pavlovian conditioning.	Salivation.	Verbal in- structions to pre- varicate.	Effect on PGR.

suspected patient is brought for diagnosis, the clinician begins to interview the patient to determine whether he is in fact mentally ill, how severe the condition is and what can be done about alleviating it. We shall consider here only the diagnostic procedure. What does it consist of?

Here again, the chief concern of the diagnostician is to determine whether the behavior exhibited by the patient is the type you would expect from a normal individual. We probe for feelings of hostility, depression, aggression, anxiety, apathy, affect, delusions, hallucinations, but our probing is simply to determine whether the patient deviates from the expected in these kinds of behaviors. The best technique thus far developed for this probing is the interview, despite its subjectivity and unstandardized character. However, it need not remain as unstandardized and unreliable as

it is now. The introduction of standardized interviewing methods of the type provided in the "Mental Status Interview and Inventory," is a first step in providing a biometric approach to assessment (Spitzer et al., 1964).

The primary design for objectifying and standardizing the interview and observational approach is best exemplified by the animal observational studies reported by Hebb and Thompson (1954) on the recognition of social attitudes in chimpanzees. Noting that the seasoned caretakers refuse more often than not to say what the motivation behind an animal's act might be unless they have had considerable experience with it, Hebb and Thompson examined carefully just how the caretakers and staff members eventually develop confidence in their understanding and prediction of a given animal's behavior.

The first important point to bear in mind is that the recognition of an attitude or mental state in an animal usually does not depend on the present behavior alone, but on the perception of the relation of the behavior to what the animal does in other circumstances. The method developed for making more accurate and valid inferences regarding the attitude of the chimpanzee consisted of three steps. First, some 36 specific acts were selected as first-order categories for tabulating observed behavior. Interpretation here was at a minimum. These elements were of the following type: putting the head against the cage wire at the observer's approach; grooming the observer's skin; beating the ground or some noisemaking surface; spitting; etc.

These raw data could not provide any basis for prediction in themselves and had to be ordered into second-order categories for grouping those elements which had roughly the same behavioral significance. These categories were: "friendly behavior," "aggression," "quasi-aggression," "avoidance," and "no response." Thus, the following elements were tabulated in the "friendly behavior" category: putting the head, shoulders, back, belly, or arm against the cage wire at the observer's approach; thrusting fingers or lips through the 2-inch wire mesh, etc.; "quasi-aggression"—walking erect with the threatening swagger of the chimpanzee about to attack, throwing sand or feces, and spitting. The second-order categories provided a more systematic way of viewing the data, but individuals with similar number of tallies in each second-order category would still exhibit quite different behavioral patterns. In order to circumvent this difficulty, a third-order category of behavior was introduced and was defined as follows: "(a) The occurrence of behavior from two second-order categories in definite sequence (e.g., avoidance followed by aggression * * * ; aggression followed by avoidance * * * ;) or (b) the (isolated) occurrence of an act in one second-

order category (e.g., aggression not preceded by quasi-aggression)."

The significance of the change from second-order to third-order categories is demonstrated dramatically in the relation of male versus female animals' behavior to the staff.

No comparison was made of the first-order categories because of their disparate nature and uneven distribution. A comparison of the second-order categories indicated that although the two sexes were equally friendly, the males excelled in quasi-aggression and in aggression. However, when the third-order categories were compared, it was found that the pattern of quasi-aggression followed by aggression occurred only in males, while seductive, friendly behavior followed by aggression occurred only in the females.

It is clear that mere atomistic enumeration of behavior is difficult to integrate into meaningful observations. However, second and third order categories and even higher categories where three or more events are included in a pattern promises to be useful. The parallel method followed in the evaluation of observed behaviors during structured interviews is to first note the molecules of behavior in items, then group the items into second-order categories or scales and finally find the pattern of scales which represent the profiles of the various groups of patients.

Reliability and Validity of Scaling

It is quite apparent that each member of the clinical team, the clinical psychiatrist, clinical psychologist, ward nurse, attendant, social worker, occupational and physical therapist, etc., makes some unique contribution to the understanding of the patient's pathology. Because of the different observational situations afforded each member, a different technique is required for eliciting and recording the observations in each clinical context. The development of such techniques was begun by our Biometric Re-

search Group and the following techniques are now available in various stages of completion: (1) The Ward Behavior Inventory, (2) Mental Status Schedule, (3) Structured Clinical Interview, (4) Children's Behavior Inventory, and (5) Social Adaptation Inventory and Anamnesis.

The Ward Behavior Inventory
(Burdock et al., 1960)

The first result of our efforts was the provision of an instrument for recording the behavior of the patient as observed by nurses, attendants, or special observers on the ward. Material relevant to ward behavior was culled from the literature and in 1953 a Hospital Adjustment Rating Scale was introduced for use in our prognostic studies at the Brooklyn State Hospital. This became the starting point for the development of the Ward Behavior Inventory.

This instrument capitalizes on the observational opportunities afforded nurses and ward attendants in everyday contact with the patients. The WBI consists of 138 items rated true or not true. All the items describe observable behaviors as seen during a specified interval of a single day. Neither retrospective material nor dynamic inferences are included. Total scores reflect global pathology and provide evidence of change in response to treatment. The Ward Behavior Inventory has proved itself reliable when raters are properly trained and motivated (Burdock, Hakerem, Hardesty, and Zubin, 1960). Reliability coefficients ranged from 0.40 for untrained raters to 0.84 for well-trained and well-motivated observers. Followup studies indicate that the global pathology score correlates at a low but significant level with outcome in terms of length of time out of hospital. WBI scores of 107 patients followed for 1 year had a correlation of 0.23 with an Outcome Index (1961). Several drug studies have proved the instrument sensitive to behavioral change. Among these are the NIMH 9-hospital collaborative study of

phenothiazine therapy conducted by the Psychopharmacology Service Center (Cole, J., 1964) and an unpublished study of the efficacy of a psychic energizer for seniles.

Future work with the WBI will focus on analysis of patterns of items and identification of clusters of patients possessing similar symptomatology. In order to obtain more objective data than those made available by the overburdened ward nurse, specially trained research nurses are to be included in our research team. The new data to be used in the construction of subscales will be based on a suitably sampled population of patients drawn from the Washington Heights area in New York City.

The Mental Status Schedule
(Spitzer et al., 1964)

In 1943 I assisted Dr. Nolan D. C. Lewis in the revision of the third edition of the "Outlines of Psychiatric Examination," the classic handbook for guiding the new resident in his mental examination of patients. A mimeographed sheet was provided for checking the pertinent items regarding the psychopathology of the patient. This was used for a short time and was included in the case history, but the rapid changes brought on by the end of World War II soon overwhelmed the training of residents and the use of the Mental Status Schedule fell into discard.

Beginning in 1960, a new attempt at standardizing the interviewing of psychiatric patients was initiated. This culminated in the development of two instruments, the Mental Status Schedule and the Structured Clinical Interview.

The Mental Status Schedule consists of an interview schedule for the mental status examination and a matching inventory of 248 dichotomous items descriptive of small units of pathological behavior. The schedule contains questions arranged in a definite sequence designed to provide for followup of incomplete responses. Most of the ques-

tions are open-ended so as to encourage the patient to reveal his own mentation. Properly administered, the interview has the feel of the clinical evaluation. However, unlike the usual clinical interview, the provision of a specific schedule of questions, a fixed order of presentation, and a uniform coverage of the same areas of psychopathology with each patient makes it more likely that the differences observed will be due to actual differences among patients rather than to different interviewing procedures. This technique has yielded reliabilities of the order of 0.90 or more when groups of patients were evaluated independently by three psychiatrists. Moreover, it has distinguished significantly between the amount of psychopathology shown by inpatients, clinic outpatients, and former inpatients on followup.

Recently, scoring systems for two types of subscales have been developed.

Clinical subscales.—Items from the MSS were grouped into 20 scales which seemed to reflect the clinical dimensions most frequently used in descriptive psychiatry. These subscales are being revised and improved by analyzing the data from a study in which clinicians made global clinical judgments on the 20 dimensions after they had interviewed each of a series of patients with the MSS. The contribution of each item in a scale to the clinical judgment of that dimension was estimated in order to determine the weight which the clinician gave to the item and thus to test whether its placement in the subscale was correct. Since some of the subscales have been revised, we plan to retest their reliability and validity.

Factor subscales.—The 248 MSS items have been grouped into 138 clusters of from 1 to 4 items each in order to have a more manageable number of variables for factor analysis. These clusters were formed according to the following principles: (1) items were combined which reflected different aspects of the same behavior or feelings; (2) items reflecting an observed

behavior were not combined with items reflecting reported behavior; and (3) items associated with different diagnostic categories were not combined. The correlation matrix for the 138 clusters of items has been obtained from a sample of 529 patients, and the process of extracting factors has begun.

Of late we have devised a series of programs for biometrics research for computing for each patient or for a group of patients (1) the total score (number of items judged true), (2) the 20 clinical subscale values (above), (3) the correlational matrix for these 20 clinical subscales, (4) the grouping and scoring of the 138 clusters (above), and (5) the individual item frequencies. These programs are available for use by investigators employing the MSS, and a computer program is being developed for the scoring of the factor subscales.

Audiotape recordings of MSS interviews have been made for training interviewers and for estimating their reliability and bias. A method for estimating accuracy has been developed, whereby a number of experienced clinicians have come to a consensus as to which items are true, false, or possible for a number of teaching tapes. A rater listens to and scores a number of recordings, and later notes for each item whether he had a "hit" (he and the experts both agree that the item is either true or false), an "add" (he rated it true, experts rate it false), or a "miss" (he rated it false, experts rate it true). The degree of inaccuracy and the direction and sources of bias are thus indicated to the rater.

Structured Clinical Interview (Burdock, 1962)

The purpose of this instrument is to focus on the general social and psychological adjustment of the interviewee. It consists of open-ended questions that are quite neutral in tone and content but which nevertheless permit the interviewee to express psychopathological ideation and

behavior. For this reason it is most suitable as an initial instrument for community surveys as well as for admission services in hospitals and clinics.

If this instrument yields a score which makes the interviewer suspect that he is dealing with a possible psychiatric case, more probing instruments like the Mental Status Interview can be applied to determine further the nature and extent of the psychopathology. If the interviewee fails to show any psychopathology on this instrument he can be tentatively regarded as normal. Interviewees who give evidence of psychopathology in this neutral instrument tend to have severer illness than those who require probing to elicit their psychopathology.

The instrument consists of an interview schedule of open-ended questions and 178 inventory items which are to be marked "true" or "not true" on the basis of the interviewee's answers. The reliability of the total score, which represents the severity of psychopathology exhibited by the interviewee, has been studied by comparing the results of two or more simultaneous but independent observers. Intraclass correlations among observers have ranged from 0.83 to 0.92 for groups of patients ranging in size from 26 to 83. With score on the Ward Behavior Inventory as a criterion for validity, the coefficients for the SCI of 0.30 to 0.68 were found, while the coefficient for an unstructured interview with the same criterion was only 0.22. The following subscales have been developed: Anger-hostility, conceptual dysfunction, fear and worry, incongruous behavior, incongruous ideation, lethargy-dejection, perceptual dysfunction, physical complaints, self-depreciation, and sexual deviance.

Future work with the SCI will be in the direction of developing norms based on suitable samples drawn from the general population and from patients.

Social Adaptation Schedule (in preparation)

This technique follows the form of the Mental Status Schedule in that the patient is examined by means of a structured interview and observations are recorded on an inventory of dichotomous items descriptive of small units of pathological behavior. However, it differs from the examination of the mental status in that the focus is not on symptomatology but on disturbance of function in the context of social adaptation. The patient is examined for evidence of disturbed functioning in the following areas of adaptation: (1) Use of leisure time, (2) friendship patterns and involvement in social activities, (3) work adjustment, (4) sexual and marital adjustment, (5) school or vocational training, and (6) level of aspiration. The technique has been designed so that it can be used to supplement the examination of the mental status or can be used on a separate occasion as an independent instrument.³

With regard to the anamnesis, we have not yet developed an instrument for this area. It is important, however, to realize that obtaining information on the entire spectrum of a person's developmental history is a research task of the first magnitude. In the experimental approach to this problem it is necessary to develop techniques which will focus on particular critical periods in the history of the individual. It is proposed to begin with a focused interview dealing with the question of the status of the patient a year before he came for help and, by focussing on this particular period, to arrive at more objective information.

A second focus might be the adolescent friendship patterns of the patient, since in some current research we have found this phase of development important in prognosis, especially with regard to type of onset. A third focus might bear on early

³ The Mental Status Schedule has recently been combined with the Social Adaptation Schedule to constitute a more comprehensive instrument, known as the Psychiatric Status Schedule.

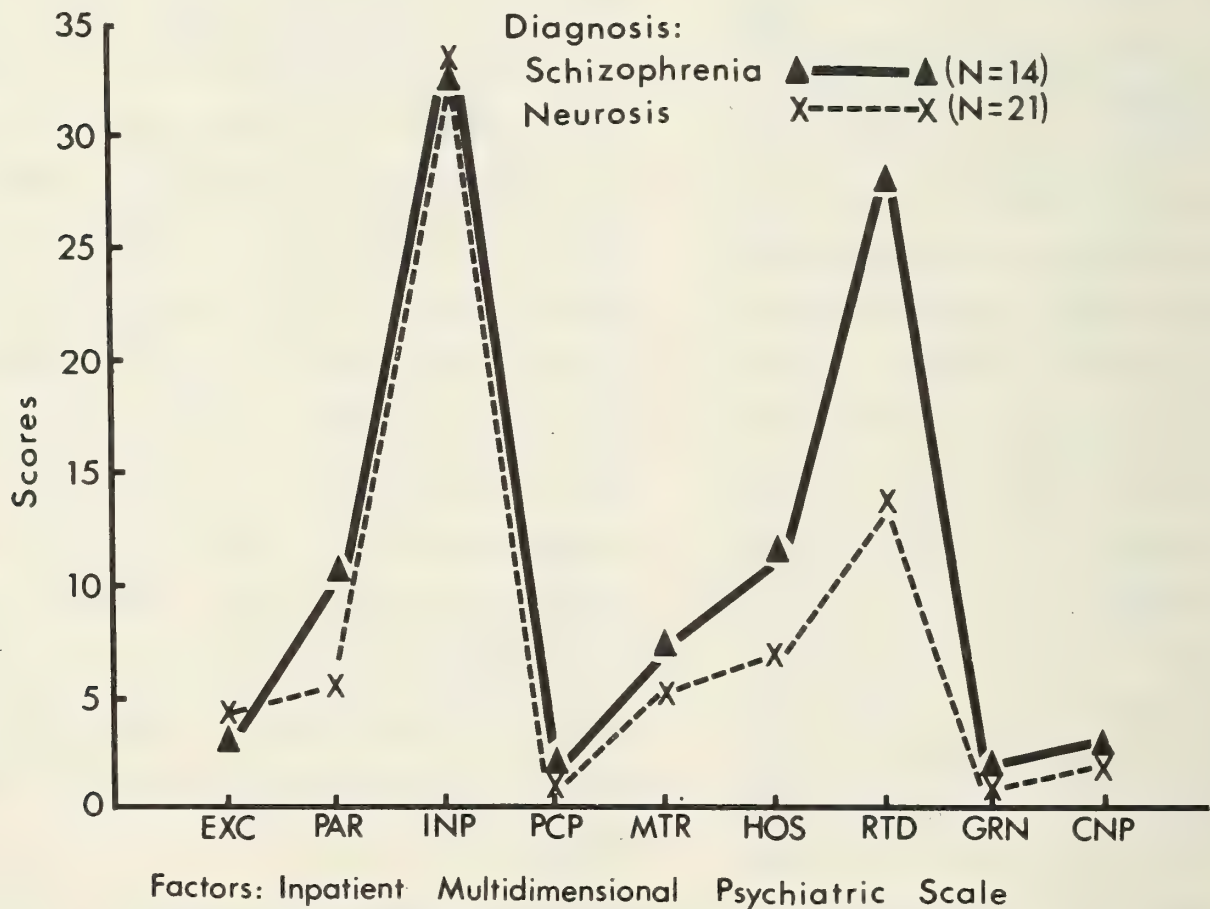
childhood, especially based on history obtained from the family. In all of these aspects of the anamnesis, at least two members of the family ought to be interviewed for checking purposes. While this may sound impractical for everyday use, it should not be beyond the range of a research project. In fact, the Katz Adjustment Scales (Katz, M., and Lyster, S. B., 1963) are based upon interviews with relatives of patients. Hopefully, the more intensive methods used in this research can form the basis for more practical approaches in the future.

The Children's Behavior Inventory
(Burdock & Hardesty, 1964)

This inventory is being used in the collection of observations of children hospital-

ized for physical illness as well as for collection of observations in children's services of psychiatric facilities. Three successive 2-hour observations are made on an observation day by specially trained psychologist observers. This technique allows for determination of frequency of pathological behaviors over the observation day. Comparisons of the hospitalized normals with the psychiatric patients will lead to development of normative criteria.

An example of the value of such systematic interviews is demonstrated by the following experiment (Katz et al., 1966). One of our structured interviews was videotaped so that it could be presented to groups of psychiatrists for their clinical judgment and diagnosis.



Mean symptom profiles of a psychiatric patient as viewed by two groups of clinicians who differed on diagnosis

(Reproduced by permission from Katz, M.M., Cole, J.O., & Lowery, Henri A. Studies of the diagnostic process. Unpublished manuscript.)

In addition to the overall diagnosis the psychiatrists were asked to fill out ratings on an inventory for such factors as excitement, paranoid projection, anxious intro-punitiveness, perceptual distortion, motor disturbances, hostile belligerence, apathy and retardation, grandiose expansiveness and thinking disorganization (Lorr, 1953). The psychiatrists were all seasoned veterans of psychiatry; nevertheless, of the 35 participants, 14 diagnosed the patient as neurotic and 21 as psychotic. An examination of their ratings, revealed however, that the groups differed significantly only in one respect—the rating on apathy. Those who rated the patient high on apathy diagnosed him as psychotic, while those who rated him low on apathy diagnosed him as neurotic.

Culture-Fair Techniques

In addition to direct interviewing with the structured instruments, certain focussed aspects of interviewing procedures have been utilized for more objective evaluation of specific patient characteristics such as flatness of affect, intelligibility of speech, and similar traits. While the structured interviews may elicit material that is largely culture dependent, the techniques to be discussed now might belong in the culture-fair category. By this we mean techniques that, though imbedded in a given culture, nevertheless have equivalents in other cultures. Examples of such culture-fair behaviors are response to greeting, bereavement, reinforcement, or praise, etc. All cultures exhibit behavior in response to such situations, but the type of behavior is colored by local norms.

In the course of our studies on the experimental analysis of the interview we have found that the rate of emission of self-referred affect statements can be lawfully modified in acute schizophrenics by the asking of questions which serve as discriminative stimuli for eliciting affect (described in Zubin, 1958), and by reinforcers

such as verbal expression of agreement—e.g., *mhmm*, *yeah*, etc. (described in Zubin, 1958; Salzinger and Pisoni, 1958, 1960, 1961). We were able to show that the process of conditioning is effective both in the beginning or middle of the interview, that a minimum number of reinforcements is necessary for conditioning to take place, and that rate of extinction varies as a function of rate of acquisition (Salzinger and Pisoni, 1961).

Rate of emission of affect statements was also successfully modified in normal subjects, and a comparison of normal and acute schizophrenic subjects showed the schizophrenics to extinguish faster than the normals (Salzinger and Pisoni, 1960).

Similar interviews with chronic schizophrenics led us to conclude that they can be more accurately described as having a low rate of responding in general, rather than as having “flatness of affect” in particular. Their low operant levels for self-referred affect made it impossible to measure their reactivity to reinforcement (Salzinger and Portnoy, 1964).

Most recently we have devised a technique which makes interviewer questions unnecessary. Under these conditions we have obtained monologs—extended uninterrupted samples of continuous speech from acute schizophrenics. We found that, depending upon the specific reinforcement contingency in effect, we could either increase general speech rate or increase the rate of a specific response class like self-referred affect (Salzinger, Portnoy, and Feldman, 1964a).

Comparisons of the rate of repetition of words in schizophrenic and normal speech samples revealed that schizophrenics tended to use the same words more frequently than normals for passages of 900 words in length (Hammer and Salzinger, 1964).

Furthermore, rate of conditioning of schizophrenic patients (i.e., the degree to which a patient becomes influenced by the interviewer) appears to be a prognostic indicator. Those patients who condition

better are more likely to leave the hospital at the end of a 180-day period (Salzinger and Portnoy, 1964).

Communicability

In order to measure the degree of communicability of schizophrenic speech, 200-word samples from the above-mentioned interviews and monologs were selected and subjected to the cloze procedure (the cloze procedure consists of systematically deleting words from continuous speech passages and having other subjects guess the words which have been so deleted). Schizophrenics' and normals' samples were compared and we found that normals communicate better than schizophrenics (a larger number of correct words were guessed). Furthermore, differentiation between the patient and normal was better for the second 100 words of the sample than for the first 100 words (Salzinger, Portnoy, and Feldman, 1964b).

We have recently replicated the above study using the same samples but a different group of predictors and separating the 100-word samples (each subject got all samples, but the 200 words were not presented consecutively). The results were almost identical with those obtained in the earlier study. Despite the separation of the two 100-word samples the second 100 words differentiated better the schizophrenic and normal than the first 100 words, and speech from schizophrenics continued to be less communicable.

We have also added two other techniques for assessing the communicability of schizophrenic speech. The first we have referred to as the "method of reconstruction." It consists of having subjects sort speech segments of equal length into the order they think they were emitted by the speaker. Comparisons of the schizophrenic and normal samples revealed that subjects had greater success in restoring the original order of the normal than of the schizo-

phrenic samples (Salzinger, Portnoy, and Feldman, 1966).

The second procedure we have referred to as the "method of unitization." It consists of having subjects divide unpunctuated speech samples into grammatical sentences and cross out words they feel do not belong. The results indicated that schizophrenic speech is characterized by having a larger number of intrusions (i.e., words crossed out) than normal speech samples. Number of units and length of unit did not differentiate normals from schizophrenics.

We have also related the cloze procedure and reconstruction method results to outcome of illness and found that the better a patient communicates the shorter his stay in the hospital, and the more difficulty the subjects found in ordering the speech samples of patients the longer these patients remained in the hospital.

The results of these more recent studies together with the results of the earlier ones have led us to propose the notion that the basic deficit in the schizophrenic patient's behavior consists of his being controlled primarily by stimuli immediate in space or in time. It is for this reason that he has difficulty in sorting, shows less object constancy (at least in the more deteriorated stages), that he extinguishes faster after conditioning, and that in language he shows "loose association" and poor ability to communicate—since language communication requires that the speaker react to long-range stimuli and not simply to short-range associations (Salzinger et al., 1966).

Culture-Free Indicators

Though no measures can be said to be completely culture free, the way in which culture might affect such measures as pupillary response to light stimuli is indirect, unlike the direct way in which culture influences primarily conceptual measures like vocabulary. The major way in which culture will tend to influence the

culture-free or -fair tests is likely to lie not in the function under measurement but in the subject's approach to the testing situation—e.g., in the subject's understanding of the purpose of the test, in the degree of fear experienced, in his motivation, attention, and cooperation, etc.—in other words, specifically in those variables which also tend to contaminate comparisons of schizophrenics and normals even when they come from the same cultural background. But it is these very contaminations that the experimental culture-fair or -free techniques have been constructed to minimize.⁴

One of the best established findings with regard to schizophrenia is that schizophrenics are retarded in reaction time as compared to normals. Since speed of response is to a large extent motivationally determined, it could readily be argued that the difference between schizophrenics and normals with regard to reaction time is not specific to schizophrenia. In order to eliminate this source of irrelevant difference, motivation must be either measured, permitted to vary at random with regard to the experimental conditions, or kept at a high level.

Since measurement of motivation is rather difficult, we chose first to sample from situations in which motivation was uncontrolled. The subject was instructed to lift his finger off a brass plate just as soon as he saw the light or heard the sound which served as stimuli. There were two light stimuli, red and green, and two sound stimuli, high and low tones. The stimuli were presented in random order. Our chief concern was not with the primary reaction times but with the influence on reaction time of shifting from one stimulus to another. There were 3 types of sequences possible: (1) No shift, the

second stimulus being identical with the first; (2) shift within the modality (red to green or vice versa, high to low tones or vice versa)—ipsimodal; and (3) shift between modalities—crossmodal (Sutton, S., and Zubin, J., 1965).

In each case, the reaction time to the second member of the pair of stimuli is recorded. We found that normals are faster on all accounts, and that reactive schizophrenics are faster than process schizophrenics. The influence of overall speed was eliminated by performing covariance analyses comparing groups on cross-modal stimuli and on ipsimodal stimuli with performance on the no-shift category held constant. Both of these analyses discriminated patients from normals on reaction to sound, but not on reactions to light. Analyses comparing process and reactive patients were not significant.

The experiment has been repeated with a number of variations in a number of populations. In summary, it may be said that there is a consistent tendency across several studies for schizophrenic males to be disproportionately impaired on reactions to sound stimuli if the stimulus in the previous trial is a light. These findings hold in both acute and chronic patients. This impairment is over and above the differences in overall level of response between groups. Comparable analyses for reactions to light do not yield differences between patients and normals.

In a comparable study done by Benton et al. (1962), on brain damaged subjects, at Iowa, there was an exactly opposite finding. The shift from light to sound did not discriminate brain-damaged subjects from normals, but the shift from sound to light did. This intriguing reversal between brain-damaged subjects and schizophrenics has not to my knowledge been followed up.

In interpreting these data one must make some assumptions about the state of readiness of the subject. It is reasonable to assume that, in all reaction time experiments, some aspect of readiness of the subject is

⁴ Such designs include obtaining complete functions rather than isolated points of measurement for each individual, comparisons of slopes of functions rather than level, the measurement of functions under idling state conditions and under load conditions, and finally, the use of the range of variation within a population as a basis for assessing observed deviation.

fluctuating from moment to moment. For these experiments, there must also be some such process, although how one prepares to observe with the eye rather than with the ear, or vice versa, is less apparent. In this connection, one might speculate on the possible role of descending inhibitory pathways which are present in most if not all sensory modalities. Shorter reaction time may then be taken as evidence that the stimulus occurred at a point in time corresponding to maximal readiness, or that the stimulus is one for which maximal readiness has been made. When the stimuli do not fall in with these expectancies, there is retardation or lengthening of reaction time. As has been shown, this state of readiness is strongly influenced by the character of the preceding trial—viz, whether the stimuli are in the same or a different sensory modality.

With respect to the differences between schizophrenic patients and normals, the studies strongly suggest that the state of readiness of the patient is disproportionately affected by events which are recent in time. In other words the patient is over-prepared by his last trial and reaction time is lengthened when the unprepared-for stimulus appears. In the random program, since all sequences are equally probable, best performance would occur if the subject assumed (correctly) that the next stimulus may be of any of the four used. Therefore preparation for any one type of stimulus is maladaptive.

Our further studies with normals have made it apparent that in these experiments there are two effects rather than one at work, one of which may be called the probability effect, in which reaction time to the uncertain or less probable stimulus is lengthened, and the other of which may be called the sensory effect. The sensory effect is revealed by the fact that manipulation of the probability of crossmodal and ipsimodal stimulus does not produce identical results—there is something special about cross-

modal stimuli which makes them peculiarly vulnerable to the state of certitude of the subject. We have not yet decided whether the differences between patients and normals are due to one of these effects, or to both. These observations have, incidentally, been supported in our evoked potential studies. In these we have found dramatic amplitude increases in the evoked potential to sensory stimuli when the subject did not know in advance which of two stimuli would be presented. However, the program in which the cross-modal stimulus is uncertain is the one that has produced the larger amplitude changes, presumably as a result of these subjects' greater degree of incertitude (Sutton et al., 1964).

Time does not permit a complete summary of our results in pupillography, nor certain other areas where cultural explanations of the observed differences are hard to find. In an initial study we found that the dark-adapted size of the pupil of acute schizophrenics, but not of chronics, is smaller than that of comparable normals (Hakerem, Sutton, and Zubin, 1964). This held true significantly for males but does not reach significance in females. Both the acute and chronic schizophrenics reach the point of maximum contraction upon stimulation by light sooner than normals. However, there were no differences between patients and normals in the extent of contraction. We concluded following the interpretation of pupillary diameter offered by Lowenstein and Loewenfeld (1950) that the acute patients show a deficiency in the inhibitory control of the Westphal-Edinger nucleus by higher centers. In order to demonstrate this deficiency in chronic schizophrenics, we have to subject the pupil to a stimulus, or put it under some load.

However, our more recent work has seemed to show almost opposite results, in the sense that normals and schizophrenics (both acute and chronic), showed similar initial pupil diameters after standard dark-adaptation but both groups of schizophrenics showed contraction amplitudes

markedly different from the normals'. In our previous experiments, the light stimulus was of one-second duration, while more recently we used briefer stimuli averaging 30 to 40 msec. A series of parametric studies will obviously be needed to resolve the apparent conflict. The nature of the patient sample must be more completely specified as to history, treatment, accuracy of diagnosis, etc., and a possible curvilinear light-stimulation function investigated. It may be that measurements close to visual threshold may yield differentiating results quite different from those obtained under supraliminal loads.

Instead of sampling or measuring motivation we can take another tack. If we find functions in which patients are more sensitive to the stimulus input than normals, we might deduce that motivation must be at a high pitch. We thought that the response to delayed auditory feedback of speech might be such a function, since schizophrenics might be bothered less by the delay and could, therefore, perform better. Unfortunately, this is not the case; schizophrenics have shown greater interference than normals. We have, however, found several functions in which schizophrenics excel. Although these results are still tentative, they do encourage the hope that more results of this type may be expected. The Bunsen-Roscoe or Bloch's law seems to hold for normals when reaction time is used as the response to the varying parameters of intensity and duration during a critical period of 10 to 20 msec. Schizophrenics apparently have not heard of this law and fail to obey it, or have a much shorter critical duration than normal.

Similarly in dichotic auditory stimulation experiments, when one ear receives either a more intense or a slightly earlier stimulus than the other ear, the localization of the sound will deviate from the center in the direction of the ear receiving the more intense or earlier click. The intensity and time differences can be matched so that the subject cannot discriminate the displace-

ment from center brought about by the two techniques. Preliminary results indicate that values cannot be found at which the schizophrenic will not discriminate the ΔI stimuli (intensity differences) from the Δt stimuli (time differences). In other words, the schizophrenic is detecting differences which the normal cannot detect. In several other ways, it becomes apparent that greater sensitivity to, or lack of integration of, input is characteristic of the schizophrenic. It is very difficult to find cultural explanations for such differences.

One of the more exciting techniques for the study of differences between patients and normals is the use of evoked potentials by Shagass and Schwartz at Iowa (1961; 1962) to measure cortical excitability and responsiveness. These workers recorded evoked potentials to two discrete, electrical stimuli presented to the wrist and separated by brief periods of time. From these electrophysiological data they plotted the resulting recovery functions for the effects of the first stimulus upon the amplitude of the evoked potential to the second stimulus. Some types of mental patients—depressives—were found to require a longer temporal separation between pulses than normal subjects for the amplitude of the second evoked potential to return to its base line, suggesting that they display less cortical excitability and responsiveness than normals.

With the aid of objective measures of both varieties—the culture-bound as well as the culture-free—a new approach can be made to the development of more homogeneous diagnostic groupings. At the present time, much of the research in psychopathology suffers from the great heterogeneity which characterizes today's diagnostic classifications. Even when there is full agreement on the part of a diagnostic team regarding such diagnoses as schizophrenia or neurosis, the individual patients in each category are far from a homogeneous group even with respect to the variables which presumably characterize each category. Bleuler

has pointed to this fact by substituting for Kraepelin's "dementia praecox," the term "group of schizophrenias." The heterogeneity in the category of neurosis is too well known to require documentation.

Heretofore, the technique applied to the statistical classification of patients has been factor analysis, either direct or inverse, of the correlations obtained from a sample. Such techniques are elegant and quite suitable if the underlying population from which the sample is drawn has a multivariate normal distribution. Were this the case, there would be no need to search for better classifications, because to begin with, our population would already be homogeneous. However, if we start with an assumption of homogeneity and wind up with a conclusion of heterogeneity, the only justifiable procedure is to reject factor analysis as inapplicable.

For this reason, once heterogeneity is established, we discard factor analysis and base our new analysis on the test profile of each patient expressed in standard scores. The usual method of correlating profiles (inverse factor analysis) takes into consideration only the shape of the profiles and neglects entirely the distance between profiles. For this reason, we deal with two measures of distance: discrepancy in shape and discrepancy in level, and group together all patients who show the greatest similarity with respect to these two measures. In this way, under certain assumptions (normality of multivariate distributions as well as zero correlation within each subgroup or some similar assumptions) a fractionation of the sample into homogeneous subgroups becomes possible.

REFERENCES:

BATESON, G., JACKSON, D. D., HALEY, J., and WEAKLAND, J. H. "Toward a Theory of Schizophrenia." *Behav. Sci.*, 1956, 1, 251-264.

BENTON, A. L., SUTTON, S., KENNEDY, J. A., and BROKAW, J. R. "The Cross-modal Retardation in Reaction Time of Patients With Cerebral Disease." *J. nerv. ment. Dis.*, 1962, 135, 413-418.

BURDOCK, E. I. "Structured Clinical Interview and Inventory." Paper presented at the Eastern Psychological Association, Atlantic City, May, 1962.

BURDOCK, E. I., HAKEREM, G., HARDESTY, ANNE S., and ZUBIN, J. "A Ward Behavior Rating Scale for Mental Hospital Patients." *J. clin. Psychol.*, 1960, 16, 246-247.

BURDOCK, E. I., and HARDESTY, ANNE. "A Children's Behavior Diagnostic Inventory." *Ann. N.Y. Acad. Sci.*, 1964, 105, 890-896.

BURDOCK, E. I., SUTTON, S., and ZUBIN, J. "Personality and Psychopathology." *J. abn. soc. Psychol.*, 1958, 56, 18-30.

COLE, J. O. "Evaluation of Drug Treatment in Psychiatry." In P. H. Hoch and J. Zubin (editors), *The Evaluation of Psychiatric Treatment*. New York: Grune & Stratton, 1964.

HAKEREM, G., SUTTON, S., and ZUBIN, J. "Pupillary Reactions to Light in Schizophrenic Patients and Normals." *Ann. N.Y. Acad. Sci.*, 1964, 105, 820-831.

HAMMER, MURIEL, and SALZINGER, K. "Some Formal Characteristics of Schizophrenic Speech as a Measure of Social Deviance." *Ann. N.Y. Acad. Sci.*, 1964, 105, 861-889.

HARLOW, H. F. "The Heterosexual Affectional System in Monkeys." *Amer. Psychol.*, 1962, 17, 1-9.

HEBB, D. O., and THOMPSON, W. R. "The Social Significance of Animal Studies." In G. Lindzey (editor), *Handbook of Social Psychology*. Cambridge, Mass.: Addison Wesley, 1954, v. I.

JASPERS, KARL. *General Psychopathology*. Chicago: Univ. of Chicago, 1963. "The Human Species," p. 667.

KATZ, M., and LYERLY, S. B. "Methods for Measuring Adjustment and Social Behavior in the Community." *Psychol Rep. Monogr., Suppl.* 4, 1953, 13, 503-535.

KATZ, M. M., COLE, J. O., and LOWERY, HENRI. "Studies on The Diagnostic Process." Paper presented at the American Psychiatric Association, Atlantic City, May 1966.

KRAEPELIN, E. "Der psychologische Versuch in der Psychiatrie." *Psychol. Arbeiten*, 1896, 1, 77.

LORR, M. "Multidimensional Scale for Rating Psychiatric Patients: Hospital Form." *Vet. Adm. Tech. Bull.*, 1953, 1, No. 10-507.

LOWENSTEIN, O., and LOEWENFELD, I. E. "Mutual Role of Sympathetic and Parasympathetic Systems in Shaping of the Pupillary Reflex to Light. Role of Sympathetic and Parasympathetic Systems in Reflex Dilation of the Pupil: Pupillographic Studies." *Arch. Neurol. Psychiat.*, 1950, 64, 341-377: 313-340.

MACHI, V. S. "An Analysis of Interperson Correlations Among Thirty Psychotics." Unpublished Ph. D. dissertation, Columbia Univ., 1956.

- MEDNICK, S. A. "A Learning Theory Approach to Research in Schizophrenia." *Psychol. Bull.*, 1958, 55, 316-327.
- MELZACK, R., and SCOTT, T. H. "The Effects of Early Experience on the Response to Pain." *J. comp. physiol. Psychol.*, 1957, 50, 155-161.
- PASAMANICK, B. "Epidemiological Investigations of Some Prenatal Factors in the Production of Neuropsychiatric Disorder." In P. H. Hoch and J. Zubin (editors), *Comparative Epidemiology of the Mental Disorders*. New York: Grune & Stratton, 1961.
- PFEFFER, MARJORIE. "Pattern of Abilities in Chronic Schizophrenics." Unpublished Ph. D. dissertation, Columbia Univ., 1952.
- SALZINGER, K., and PISONI, STEPHANIE. "Reinforcement of Affect Responses of Schizophrenics During the Clinical Interview." *J. abn. soc. Psychol.*, 1958, 57, 84-90.
- SALZINGER, K., and PISONI, STEPHANIE. "Reinforcement of Verbal Affect Responses of Normal Subjects During the Interview." *J. abn. soc. Psychol.*, 1960, 60, 127-130.
- SALZINGER, K., and PISONI, STEPHANIE. "Some Parameters of the Conditioning of Verbal Affect Responses in Schizophrenic Subjects." *J. abn. soc. Psychol.*, 1961, 63, 511-516.
- SALZINGER, K., and PORTNOY, STEPHANIE. "The Conditionability of Different Verbal Response Classes: Positive, Negative and Non-affect Statements." *J. gen. Psychol.*, 1964, 70, 311-323.
- SALZINGER, K., PORTNOY, STEPHANIE, and FELDMAN, R. S. "Experimental Manipulation of Continuous Speech in Schizophrenic Patients." *J. abn. soc. Psychol.*, 1964, 68, 508-516 (a).
- SALZINGER, K., PORTNOY, STEPHANIE, and FELDMAN, R. S. "Verbal Behavior of Schizophrenic and Normal Subjects." *Ann. N.Y. Acad. Sci.*, 1964, 105, 845-860 (b).
- SALZINGER, K., PORTNOY, STEPHANIE, and FELDMAN, R. S. "Verbal Behavior in Schizophrenics and Some Comments Toward a Theory of Schizophrenia." In P. H. Hoch and J. Zubin (editors), *Psychopathology of Schizophrenia*. New York: Grune & Stratton, 1966.
- SCHAEFER, T., JR., WEINGARTEN, F. S., and TOWNE, J. C. "Temperature Change: the Basic Variable in the Early Handling Phenomenon?" *Science*, 1962, 135, 41-42.
- SHAGASS, C. "A Neurophysiological Approach to Perceptual Psychopathology." In P. H. Hoch and J. Zubin (editors), *Psychopathology of Perception*. New York: Grune & Stratton, 1965.
- SPITZER, R. L., FLEISS, J. L., BURDOCK, E. I., and HARDESTY, ANNE S. "The Mental Status Schedule: Rationale, Reliability and Validity." *Compreh. Psychiat.*, 1964, 5, 384-395.
- SUTTON, S., BRAREN, MARGERY, JOHN, E. R., and ZUBIN, J. "Evoked Potential Correlates of Stimulus Uncertainty." *Science*, 1965, 150, 1187-88.
- SUTTON, S., BRAREN, MARGERY, and ZUBIN, J. "Sensory, Conceptual, and Emotional Components of the Evoked Response to Sound and Light Stimuli in Man." *Psychonomic Society*, Chicago, Ill., October 1965 (mimeographed).
- SUTTON, S., and ZUBIN, J. "Effect of Sequence on Reaction Time." In J. E. Birren and A. T. Welford (editors), *Behavior, Aging, and the Nervous System*. Springfield, Ill.: Charles C. Thomas, 1965.
- THOMPSON, W. R., WATSON, J. and CHARLSWORTH, W. R. "The Effects of Prenatal Maternal Stress on Offspring Behavior in Rats." *Psychol. Monogr.*, 1962, 76, 1-26.
- ZUBIN, J. "A Biometric Model for Psychopathology." In R. A. Patton (editor), *Current Trends in the Description and Analysis of Behavior*. Pittsburgh: Univ. of Pittsburgh Press, 1958.
- ZUBIN, J. "Behavioral Concomitants of the Mental Disorders—a Biometric View." In B. Wigdor (editor), *Recent Advances in the Study of Behavior Change*. Montreal: McGill Univ., 1963.
- ZUBIN, J., HAMWI, V., and WINDLE, C. "Retrospective Evaluation of Psychological Tests as Prognostic Instruments in Mental Disorders." *J. Pers.*, 1953, 21, 342-355.
- ZUBIN, J., and KATZ, M. M. "Psychopharmacology and Personality." In D. Byrne and P. Worchel (editors), *Personality Change*. New York: John Wiley, 1964.

Summary Presented at Conference:

My presentation can be condensed into the following four points: (1) The contrast between the biometric and the clinical approach of diagnosis; (2) the evaluation of the current status of diagnosis; (3) the preparation of instruments to improve present-day diagnosis; and (4) organization of the data resulting from such an attempt.

The clinical approach is too well known to require description. The essential element of a good clinical diagnosis is that it constitutes a creative act which belongs to the realm of discovery, using Reichenbach's term, rather than to the realm of verification. Such verification as is in-

dulged in consists of impressionistic testing of successive hypotheses that arise in the course of the clinical examination.

The biometric approach takes as its point of departure the objective measurement of the behavior of the patient, applies taxonomic principles for classifying such behavior in profiles of the systems, and finally appeals to statistical evaluation for determining the reliability and validity of a diagnostic label placed on the patient, as well as his prognosis and the outcome of followup.

Both approaches are essential, since, without the clinician there would be nothing to measure, and without measurements the clinician would soon become a prey to whatever wind of fashion blows.

One might say that the clinician provides the hunches and the inspiration. He is the one who storms the beachhead, while the biometrician may be likened to the foot soldier who secures the beachhead and organizes the territory, bringing law and order into its interior.

In short, without the clinician there would be no problem to solve; without the biometrician there would be no reliable solution.

There is one more difference. The final diagnosis—at least at the present stage of our development—must perforce be made by the clinician who, like the judge in court, despite the paucity of evidence, must arrive at a decision. That is why a study of the decision process itself, as Meehl and as Hammond propose, is so essential.

Biometrics is essentially nomothetic, not concerned with the diagnosis of an individual case; but it can buttress the idiographic approach very nicely.

At the present time our diagnostic schema consists of a mixture of descriptive terms observed by the clinician and etiologic considerations arising from theoretical structures. The symptoms which the patient complains of, and the signs of deviant behavior which the clinician observes, are

embroidered into a structure of postulated causes interwoven with dynamics, and the resulting structure is one that is so complex that the reliability and validity of the end results are very difficult to establish.

The symptoms and signs which constitute the present-day warp and woof of diagnostic procedure have developed over a long period of time, and have been primarily derived from the observations of patients who were hospitalized. Only now we are freeing ourselves from too great attention to "hospitalism," which was an essential feature of institutionalized mental patients, to observe behaviors which occur in the family and in the community as the mental patient begins to be seen more frequently outside of the hospital.

Because of the varying composition of the population under observation it becomes quite clear that one must understand the social-cultural milieu of the individual before the detection and the diagnosis of illness, which represents deviation and expectancy, can be made on some sound grounds.

Regarding etiology of the illness, a variety of scientific models have been provided for structuring the elements that go into possible causation. Among these are the social-cultural model, the developmental model, the learning and conditioning model, the genetic model, the environmental model, the neurophysiological model. It would take too long to go into the details of each of these; but, for the most part, it is quite clear that for any illness each of these models plays a varying role, and that we need a supermodel, such as is afforded by the epidemiological model, for weighting the different models and their proper contribution to the causation of an illness.

While waiting for the doctor—that is, for science to provide keener tools for evaluating these models—we have to resort to the best methods we now have. Tools are being developed along the following three basic lines: culture-dependent methods, culture-

fair methods, and culture-free methods. A culture-dependent method tries to determine to what extent a person's behavior deviates from the framework of the social-cultural norm. The best method now available for such purposes is the interview.

Instead of the usual free-wheeling clinical interview we have introduced systematic structured interviews yielding total scores relating to overall psychopathology, with subscales related to the individual dimensions that contribute to the total score.

These interviews cover the general area of the mental status of the patient and his performance in the following areas of life: travel, eating, personal hygiene, living arrangements, drug and alcohol usage, sex, illegal acts, leisure time, interpersonal relations, and self-injury. The role that an individual occupies in society as a mate, breadwinner, housekeeper, parent, or student, are also covered. This is included in the psychiatric status schedule which is being developed in biometric research by Dr. Spitzer and his colleagues, including Mr. Fleiss, who is here tonight.

We are also developing a psychiatric history schedule, which is designed to enable an investigator systematically to record those aspects of the subject's history considered of importance in diagnosis and prognosis. Information is obtained about the development of the present illness, nature and number of past psychiatric episodes, academic and occupational performance, adolescent and adult social and sexual adjustment, and severity and temporal characteristics of a wide range of pathological signs and syndromes.

The culture-fair method attempts to deal with those behaviors which, though rooted in a particular culture, nevertheless can be translated from one culture to another. For example, such behavior as bereavement has equivalence in almost all cultures. Greeting behavior is another example.

The only culture-fair technique that we have time to describe at the moment is the technique for assessing communicability of speech. In this method the cloze technique is used for detecting whether a given person's speech is intelligible; for example, when every fifth word is eliminated is it still possible to understand what he is trying to say? The method involves the filling in of the gaps with the word thought to be originally present, and the proportion of correctly filled gaps is a measure of the intelligibility. This technique distinguishes between schizophrenics and normals and has prognostic value.

In the area of culture-free indicators, the responses that are elicited occur during the first 1,000 milliseconds following stimulation. They occur so quickly that we think culture doesn't have a chance to get its licks in.

The primary difficulty that one faces in this area is that of ensuring the motivation of the individual, since it is well known that mentally ill people are not as well motivated as normals. Hence a technique which shows the mental patient to be poor in performance may simply reflect his lack of motivation.

There are three ways of dealing with this problem. First, measuring motivation directly, which turns out to be very difficult, as you know; second, randomizing motivation; and, third, ensuring that motivation is on a high level. The last two methods have been attempted. And, as an example of the third method—ensuring a high level of motivation—we have used techniques such as dichotic auditory stimulation, in which the patient can distinguish sensory inputs which the normal fails to distinguish. You present two sounds to both ears simultaneously and ask the person to indicate whether the sound is located in the mid-plane. If one sound reaches the ear a little earlier, or is more intense, the apparent sound localization will move in the direc-

tion of the earlier ear, or of the more intense stimulus.

We can map around the head an area where discrepancies in time and in intensity are equivalent: you can't tell the difference, if you are a normal; but the patient can tell the difference.

Another technique which we use in this general area is that of pupillography, which we can't go into in greater detail. There are several other techniques of this type, and the further discovery that the evoked potential can serve to differentiate between patients and normals gives us a method which may be relatively free from cultural influence.

We come now to the last point: organization of the data resulting from such research. Here is perhaps the primary problem facing us in this conference.

An important thing to remember is that the organization of data must always follow some goal, as Dr. Kramer sagaciously insists. The goals which we would like to set for the analysis of the data are: first, the evaluation of the severity of psychopathology present, and second, its prognosis.

For the time being we are satisfied with using concurrent validation as a measure of severity. By having a trained clinician estimate the severity of the illness we can utilize his judgment as a first approximation in the iterative process of evaluating the scores we obtain for total psychopathology as well as on the subscales. By a process of successive approximations we can perhaps eventually obtain objective measures which may be better than the clinical evaluation, in the same way that Binet's tests were an improvement on teachers' judgment of intelligence, even though the latter served as the initial criterion for intelligence.

With regard to prognosis, which is essentially concerned with predictive validity, we plan to use followup of patients to see how well they eventually turn out, and, in

this way, to develop methods for evaluating the culture-dependent, culture-fair, and the culture-free techniques.

Whether we need to turn to typological approaches, or be satisfied with dimensional approaches, will remain an open question until the data are collected and some of the statistical tests which are necessary for the detection of discontinuities in the data become available.

The presence of discontinuities is a very important matter. One often wonders whether the discontinuity inheres in the behavior of the patient or in the behavior of society. It may depend on the attitudes of the community, or on the examiner, rather than on the patient's behavior.

Furthermore, discontinuity may exist, not in a given variable but in interrelations between several variables. However, only the development of better methods or criteria for classification can determine what path to take in the future.

I have great faith—perhaps unfounded—that the master mathematician of the future can find transformations that will eliminate discontinuities, typologies, and what-not. But the fact that at the present state of our knowledge some typological analyses seem to be better than dimensional analyses, as several of the empirical papers have shown, is sufficient warrant for their usefulness.

Furthermore, disease entities, if they exist, are more "*gemütlich*" to typological than to dimensional approaches. In other words, the dimensional approach seems to go contrary to the kind of things that clinicians, in search for individual similarities, often tend to look for.

So, in summary, by providing more objective tools, better classification can ensue. Since the clinician is more concerned with individual similarities than with individual differences, typology may be a useful technique to use in assisting clinical diagnosis.

Patterns of Intellectual Functioning and Their Implications for the Dynamics of Behavior

David R. Saunders, Ph.D.¹ and John W. Gittinger,² M.S.

Introduction

If we are to attach meaning to patterns or profiles of abilities, we must assume that real individual differences in these profiles exist and that they can be observed. For a typical individual, this implies the existence of areas of relative strength and areas of relative weakness. The major psychological premise of this paper is that most individual behavior may be regarded as helping to *minimize for the individual* the significance of his weaknesses. Often, this may be accomplished by exploitation of his strengths, but other mechanisms may also be discerned. We hold this premise to be self-evident, and fundamentally important. It is important because it provides a link between the fields of "intelligence" and "personality" as they are usually conceived. Indeed, if it is possible to spell out the specific behaviors that can provide the indicated minimization, given a specific configuration of abilities, then it is possible to infer "personality" from a battery of intelligence tests. Moreover, in principle if not yet in practice, it becomes possible to define stress for an individual as any situation that interferes with his attained minimization, and to predict the direction of his behavioral efforts to re-establish it. The implications of such for education, psychotherapy, etc., are again obvious.

The minor premise from which this paper proceeds is that the Wechsler battery of subtests (Wechsler, 1944, 1955) provides valid information about individuals that

can be employed to spell out certain of the aforementioned specifics. Clearly, the Wechsler is not the only possible source for this information, nor is it optimally designed to provide all the information that may be required in every case. We regard it merely as an adequate start in the right direction. In addition, however, it is the instrument that has been most widely employed by clinicians in their studies of patterns of ability, so that both data and hypotheses are relatively easy to obtain. This study has benefited enormously from a collection of over 12,000 Wechsler profiles, gathered from sources too numerous to acknowledge here, and has benefited crucially from hypotheses implicit in the Personality Assessment System (Gittinger, 1964).

Combination of the two specified premises leads to the conclusion that we propose to consider the Wechsler seriously as a personality measure, as well as as a source of varied patterns of intellectual functioning. Actually, the Wechsler is a better instrument for this purpose than measurement specialists would have believed a few years ago, and we may begin by briefly reviewing its salient properties. We must also provide a preliminary exposition of certain major features of the Personality Assessment System (PAS), a task that will be made relatively easier by illustrative references to the Wechsler. Given this groundwork, we will direct attention to results that have been obtained using several different statistical methodologies. These results appear to be generally consistent with our premises and with the PAS. On the other hand, they may suggest inadequacies in the available methodologies for systematic analysis of intellectual pattern data.

Nature of the Wechsler

The standard scoring of the standard Wechsler provides a separate measure for

¹ University of Colorado.

² Psychological Assessment Associates, Washington, D.C.

each of 11 subtests—Information (I), Comprehension (C), Digit Span (D), Arithmetic (A), Similarities (S), Vocabulary (V), Picture Arrangement (PA), Picture Completion (PC), Block Designs (BD), Object Assembly (OA), and Digit-symbol Substitution (DS). The WISC also includes a Mazes subtest (MZ). Normally, these subtest scores are simply combined into an index of verbal ability, which amounts to an average of I, C, D, A, S, and V, an index of performance ability, which amounts to an average of PA, PC, BD, OA, and DS, and perhaps MZ, and an overall index of intelligence known as “Full-Scale IQ.” Every WAIS record form bears the following admonition from the publisher: “Clinicians who wish to draw a ‘psychograph’ on the above table may do so by connecting the subject’s raw scores. The interpretation of any such profile, however, should take into account the reliabilities of the subtests and the lower reliabilities of *differences between* subtest scores.” [Italics in the original.] This warning is probably based not only on McNemar’s disquieting remarks (1957) regarding the reliability of difference scores, but also on the results of factor analyses by several investigators that found only three to five common factors in the whole battery (Balinsky, 1941; Birren, 1952; Cohen, 1952, 1957, 1959; Gault, 1954; Goldfarb, 1941; Hagen, 1952; Hammer, 1950; Hover, 1951; Simkin, 1951). Unfortunately, both lines of argument are misleading.

The isolation of only three or four or even five common factors from the Wechsler subtest battery is an artifact of inadequate experimental design. None of the studies just cited used enough variables, according to the criteria outlined by Thurstone (1947, p. 293), to permit the definition of additional factors. Two published (Davis, 1956; Saunders, 1959) and various unpublished studies which are not subject to this artificial restriction all suggest that the Wechsler battery contains at

least as many common factors as it has subtests. In fact, factor analyses of selected sets of items (Saunders, 1960a, 1960b, 1962) has indicated that several of the subtests themselves are definitely nonhomogeneous, and that approximately 18 factors would be required to account for a complete matrix of item intercorrelations. Both the number and nature of these factors is confirmed (Saunders, 1964a) even by an analysis of items from a Japanese-language adaptation of the WAIS (Kodama et al., 1958). It is no longer possible to entertain a few-dimensional concept of this battery.

These results should serve also to undermine any further overcautious reasoning based on reliability considerations. All conventional internally-derived estimates of reliability depend on the assumption of item homogeneity, i.e., that the items involve no more than one common factor. Ordinarily, the possibility that this assumption is not met is dismissed with the observation that this can only lead to “conservative” underestimates of true reliability. The possible magnitude of underestimation error, let alone its subsequent effect on estimated reliabilities of difference scores, seems not to have been considered. For example, a 15-item subtest whose items are drawn equally from three uncorrelated but separately perfectly reliable (homogeneous) pools would have an estimated reliability of about 0.96, and the difference between two such subtests would have an estimated reliability of about 0.92. Evidently, using conservative statistics, it is not difficult to make even good difference scores look bad.

In fact, in view of the factorial evidence, the derivation of only one score from each Wechsler subtest is itself a relatively conservative procedure! Nevertheless, much of the material that it is of interest to cite here is based on the subtest structure rather than on the factor structure. This is mainly because factor scoring has not yet been applied to enough cases

to define adequate norms. Moreover, the reliability of measurement of a few of the factors really does leave something to be desired. Consequently, we shall proceed using either the subtest structure or the factor structure of the Wechsler, or both, whichever may be most feasible in the particular context. In any case, it is our thesis that these results will be found to apply in all essential respects not only to the factor structure of the Wechsler, but also in the broader context of studies like those of Guilford (1956) and Cattell (1957).

Relative Primitivity of the Wechsler Measures

We have examined, at least on a pilot basis, eight different ways of assessing the relative primitivity of the Wechsler measures. These may be indicated as follows:

- (1) Twin studies.
- (2) Studies of relative stability over time (retest reliabilities).
- (3) Factor analyses of various age cohorts.
- (4) Distribution parameters of deviations from profile average levels.
- (5) Heteroscedasticities of joint distributions.
- (6) Communalities of measures with factor analyses of known primitive measures.
- (7) Evidences for localization of function (versus mass action).
- (8) Recognition of specific physiological correlates or mechanisms.

Insofar as we have explored them, all these have tended to confirm each other (Saunders, 1963). However, at the present time, only method four has provided results based on a large enough sample to merit independent attention.

Essentially, method four provides a way of separating abilities into aptitudes and achievements. After O'Connor (1928), an aptitude is an ability whose development is controlled by maturation, and later by

deterioration, primarily beyond the control of the individual or his environment. In contrast, an achievement is an ability whose development depends on exertions of either the individual or his environment or both. Aptitudes are presumed to depend on polygenic mechanisms leading to normal distributions around average ability, in which each individual occupies his predetermined place. Achievements may be modified, however, and we may recognize individuals as underachievers, normal achievers, or overachievers. Varying degrees of underachievement are common and relatively simple to comprehend. Overachievement is a less common occurrence, particularly if we reserve the term to refer to achievement accomplished at the expense of *disproportionate* underachievement in other areas. (As we shall see below, despite its relative infrequency, the recognition of this kind of overachievement is fundamentally important in understanding the dynamics of certain patterns of ability.) This leads to the interesting prediction that distributions of deviations of achievements from their corresponding intra-individual overage abilities³ will be negatively skewed ($\beta_1 < 0$), while similar distributions for aptitudes will be symmetric ($\beta_1 = 0$). Also, by a parallel argument, the distributions of deviations of achievements should exhibit excessive kurtosis ($\beta_2 > 3$), while those for aptitudes should display normal kurtosis ($\beta_2 = 3$).

Figures bearing on these predictions have recently been developed from the large collection of Wechslers, with the following results:⁴

³ Unfortunately, none of the conventional statistical measures of central tendency provides a satisfactory definition of this intra-individual average, which occupies a central theoretical position in the PAS and is referred to as the Normal Level (NL).

⁴ These results were obtained using NL₂₇. Essentially similar results have been obtained from several runs, using successively more data and variations in definition of the Normal Level. The specific figures are subject to revision as a function of the discovery of errors in the underlying data.

	N	μ	σ	β_1	β_2
PC.....	12111	-1.99	2.10	-.177	4.380*
S.....	12111	-.98	2.02	-.427*	3.914*
I.....	12111	-.75	1.75	-.428*	3.822*
C.....	12111	-.68	2.04	-.355*	3.712*
OA.....	12111	-2.14	2.57	-.362*	3.511*
A.....	12111	-1.68	2.57	-.396*	3.402*
PA.....	12111	-2.47	2.56	-.039	3.245
DS.....	12111	-2.24	2.37	-.132	3.224
BD.....	12111	-1.69	2.24	-.207	3.131
D.....	12111	-2.30	2.95	-.030	2.943

In this table the subtests have been listed in order of *decreasing* values for β_2 , so that achievement scores are ranked above aptitude scores. The more extreme values of β_1 and β_2 are marked by asterisks; because of the very large sample, all but two values of β_1 and two values of β_2 differ from the expected normal curve values at the 0.01 level of statistical significance. It may be observed that the data provide closely parallel indications from both β_1 and β_2 . PC, with its small β_1 and relatively largest β_2 , may be regarded as an achievement that is almost as often overachievement as underachievement.

The results obtained from smaller scale analyses, using the various other methods of assessing relative primitivity, have been in good agreement with these, considering the sample sizes involved. Moreover, as will become immediately apparent, these specific results are in essentially perfect agreement with the tenets of the PAS.

Fundamentals of the Personality Assessment System (PAS)

The Personality Assessment System is rather like the proverbial elephant, having many aspects that compete for primacy of attention, each of which provides only a partial and possibly misleading concept of the whole. Perhaps there is no uniformly best way to introduce the PAS—one that will not raise more questions than it answers. (And perhaps this quality represents one of PAS's more important assets.)

In fact, however, the PAS has already

been introduced, via the major premise of this paper. The PAS is equally at home no matter whether we are discussing patterns of intellectual functioning or patterns of personality. In one field we may speak of the dependence of "achievement" on "aptitude," or in the other of the dependence of "character" on "temperament." This analogy is regarded by the PAS as a true duality, subject to substantial amplification and extension. In addition, we have already introduced evidence that some of the qualities with which the PAS must deal are more primitive than others. PAS actually postulates three levels of adjustment—a "primitive" level, an "attained," habitual or basic level, and a relatively superficial, "ideal" or contact level.

For any one individual, the primitive level and its implied behavioral tendencies are treated as given, though these may in principle be altered by organic processes or intervention. The initial developmental task of the individual is to "learn" to behave in accordance with the expectations of his environment, as these are communicated to him by his parents, peers, and teachers. This task may prove easy or difficult, depending on the "amount" of "primitive" behavior he displays, its concordance with expectations, and the kind and degree of pressure toward conformity elicited in response. The individual may discover that his primitively-determined behavior is generally acceptable, in which case his "learning" reduces to simple maturation. More typically, he will discover that his primitively-determined behavior is in some sense unacceptable, and that his overt behavior must depend at least in part upon contrary habits, as if his aptitude pattern was different from what it really is.

The acquisition of such habits, that counteract primitively-determined behavior, is referred to in PAS terms as "compensation," and the habits themselves are called "compensators." Since the underlying primitive aptitudes and correspond-

ing temperamental tendencies cannot actually be reversed, compensation produces intrapsychic tension, and we must eventually consider the capacity of the individual to deal with this byproduct. Attainment of compensation is a form of achievement, which may be brought about in degrees suggested by the terms “under-compensation,” “normal compensation,” and “overcompensation.” Most actual development of compensation is presumed to take place early in the life of the individual, independently of either awareness or volition.

It is no coincidence that the onset of adolescence coincides with the age of reason. It is at this point that awareness and volition on the part of the individual begin to modify the course of his further development, awkwardly at first, but with gradually increasing effect. It is at this point, also, that Strong’s measure of Interest Maturity (1943) leaves its baseline, and various self-report measures of personality begin to acquire reliability. To the extent that the individual creates a new and independent image of himself, he takes on the task of living up to it; we must be concerned with the consistency with which this is accomplished, and the rate at which such behavior produces habit strength commensurable with compensation. Modification of behavior brought about by this mechanism may either reinforce or detract from or substitute for the changes in behavior brought about through compensation proper, leading in each case to characteristic patterns of test performance.

A second major feature of the PAS is the notion that these three levels of adjustment can be independently recognized in several different psychological areas. The “standard” PAS (Gittinger, 1964) postulates three such areas, and labels each in terms of a pair of polar descriptors, as follows:

- E-I : ExternalizerInternalizer
- F-R : FlexibleRegulated
- U-A : Role-UnadaptableRole-Adaptable

Inasmuch as three statistically independent measures are required for the unambiguous placement of an individual within each of these areas, subtest scoring of the Wechsler does not permit utilization of any more areas than those just indicated. The subtests have been interpreted in accordance with the following scheme:

Theoretical Dimension	Primitive Measure	Compensator	Modifier
E-I F-R U-A	D BD PA	A S PC	I C OA

The standard PAS assigns a special role to the DS subtest, and discards V altogether. We note that this scheme could not be in better agreement with the subsequently obtained findings from skewness and kurtosis, reported above.

An “extended” PAS may be described if we choose to employ factorial scoring of the Wechsler. Eighteen factors appear to be just sufficient to provide unambiguous placement of individuals according to six areas of measurement, including the three standard areas. Giving consideration to various kinds of information available regarding the factors, the following scheme currently seems most reasonable:

Theoretical Dimension	Primitive Measure	Compensator	Modifier
E-I	DS D V	MZ A C ₀	OA I ₂ I ₅
F-R	BD S ₁	S ₂ PC ₂	S ₃ I ₄
U-A	PA	PC ₁	I ₃

The factors in this scheme have been designated according to the subtest with which they are most closely aligned. (Factor I₁ does not appear in the scheme because it is regarded as equivalent to a general,

second-order factor, i.e., Normal Level.) We observe, among other things, that this scheme can be reduced to the standard scheme as readily as to any other of like complexity, when the constituent factors of each subtest are recombined.

It is not essential to the reasoning of this paper that we accept the details of either of these schema as final or correct, and we therefore postpone specification of supporting arguments to another context. It is desirable, however, to assign descriptive labels—"names"—to the factors, and by extrapolation to the subtests, as a means of facilitating thinking and conversation. The following listing may serve these purposes:

A_o = Residue of Arithmetic Subtest, regarded as a measure of "Ideational Discipline." (Saunders, 1960b.)

BD = Block Design Subtest, regarded as measuring a generalized "Perceptual Threshold."

C_o = Residue of Comprehension Subtest, regarded as measuring acceptance of meaningfulness of the items and indirectly strength of Identification as a civilized human. (Saunders, 1962.)

D = Digit Span Subtest, regarded as measuring the strength of the primitive tendency toward "Internalization." Possibly related to alpha-frequency and cortical arousal via the Ascending Reticular Formation. (Saunders, 1960c, 1961.)

DS = Digit Symbol Subtest, regarded as a measure of Psychometabolic Rate. May be influenced by drugs, brain damage, or psychological state.

I_1 = General (Nonspecific) Information, regarded as a measure of total Mental Capacity. (Saunders, 1960b.)

I_2 = Current Information. (Saunders, 1960b.)

I_3 = Cross-cultural Information. (Saunders, 1960b, 1964a.)

I_4 = Generalized (Scientific) Information. (Saunders, 1960b.)

I_5 = Role-oriented (Intracultural) Information. (Saunders, 1960b.)

MZ = Mazes Subtest, regarded as a measure of "Synaptic Persistence." (Smith and Carrigan, 1959.)

OA = Object Assembly Subtest, regarded as a measure of the "Effect of Uncertainty." (Lanfeld and Saunders, 1961.)

PA = Picture Arrangement Subtest, regarded as a measure of "Conditionability." (Haronian, 1962.)

PC_1 = Maintenance of Contact, as scored from certain Picture Completion items. (Saunders, 1960a.)

PC_2 = Maintenance of Perspective, as scored from certain Picture Completion items. (Saunders, 1960a.)

S_1 = Precision of Judgment, as scored from certain Similarities items. (Saunders, 1962.)

S_2 = Emotional Control, as scored from certain Similarities items. (Saunders, 1962.)

S_3 = Psychological-mindedness or Psychological Information, as scored from certain Similarities and Comprehension items. (Saunders, 1962.)

V = Vocabulary Subtest, regarded as a measure of Symbolic Range.

It is essential to the reasoning of this paper that we appreciate the diversity of ways in which the PAS interpretation of a given performance may be influenced by the context of other performances in which it occurs. The examples in the next section are intended to illustrate these phenomena in terms of the standard PAS.

Some Illustrative PAS Interpretations

The best method of demonstrating the complexity of meaning and the sensitivity ascribed by the PAS to profile variations is by a consideration of real examples. The five examples to be presented have been selected out from the large collection because of their close and simple relation to one another, which will permit an examination of the effect of one variable at a time within a constant context provided by the patterning of certain theoretically more fundamental variables. Many other such examples could have been found, even though much more complex relationships are commonly dealt with by the system.

Case	Age	Subtest scores in WTS metric									
		D	A	I	BD	S	C	PA	PC	OA	DS
JMT 52430.....	32	6	6	13	12	10	13	9	13	13	7
JMT 54284.....	30	9	10	14	14	14	13	7	14	14	11
RAP 01017.....	31	10	10	14	14	9	15	9	10	8	8
JC 01027.....	25	6	7	14	16	12	15	9	14	15	16
JMT 59028.....	28	9	17	14	15	12	15	9	14	16	10

These are all men who have completed their college education. Three of them were tested as applicants for career positions in a management training program of a government agency. One (JC 1027) was under treatment in a VA-related institution with a diagnosis of "severe obsessive-compulsive reaction." The other (RAP 1017) is drawn from Rapaport's classic study (1945), where he carried a diagnosis of "acute unclassified schizophrenia," and

displayed the specific symptomatology of "grandiose, mystical delusions." Either in terms of the Wechsler Full-Scale IQ or the PAS Normal Level, the mental levels of all five men are comparable.

In order to derive PAS meanings from the test profiles, the deviation of each score from its respective Normal Level is determined, as follows (deviations of +1 or -1 are considered insignificant and shown as 0):

Case	NL	Subtest deviation scores									
		D	A	I	BD	S	C	PA	PC	CA	DS
JMT 52430.....	13	-7	-7	0	0	-3	0	-4	0	0	-6
JMT 54284.....	14	-5	-4	0	0	0	0	-7	0	0	-3
RAP 01017.....	14	-4	-4	0	0	-5	0	-5	-4	-6	-6
JC 01027.....	14	-8	-7	0	+2	-2	0	-5	0	0	+2
JMT 59028.....	15	-6	+2	0	0	-2	0	-6	0	0	-5

Within this sample of 5 cases, the deviations are held relatively constant for 5 of the 10 subtests—D, I, BD, C, and PA. In particular, Digit Span is always at least four points below the Normal Level, Block Design is at or above the Normal Level, and Picture Arrangement is at least four points below the Normal Level. The combination -D +BD -PA, which is usually referred to in PAS terminology as "ERU," represents one of the eight possible PAS primitive patterns. Essentially, then, all five of these cases have a dynamically common original or fundamental psychological orientation. At the risk of great oversimplification, the following—out of context (Gittinger, 1964)—is a PAS description of ERU characteristics:

The ERU child begins life as an externally-oriented, imitative but per-

severating, socially inept individual. Regardless of his Normal Level, his maturation rate will be considered erratic, for even though he learns behavioral activities quickly, he appears inefficient and somewhat awkward. He will demand considerable attention but will receive at worst rejection and humiliation and at best attention deriving more from a sense of responsibility or pity than from genuine affection and warmth. He will, in return, either learn to curb his demands for attention or else develop indirect ways of receiving it. He will be relatively vulnerable to many early childhood trauma. For example, separation from his mother could be relatively traumatic because the chances are greater that he will move from a comfortable environment in which the involvement stems from a deep sense of responsibility to one which might be

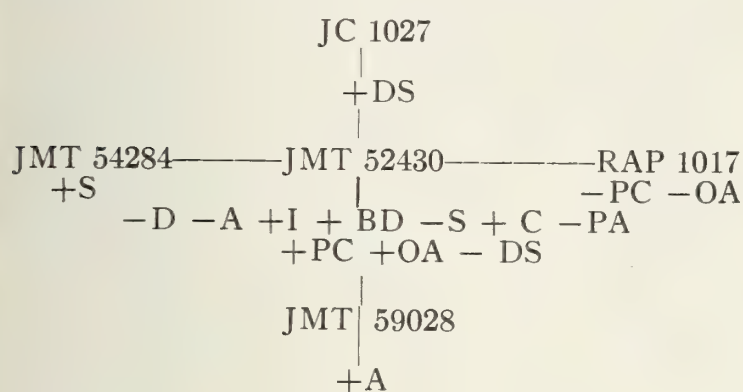
either hostile or indifferent. Although he has the ability to make rapid and effective emotional involvements with others, the chances of genuine reciprocation are not good . . . Acquisition of language would be rapid, for initially, he would have considerable need to communicate with others. This, coupled with his need to be understood, might actually result in more rapid acquisition of language by the ERU than by any other primitive group. Depending on his intelligence, he would not only learn vocabulary or languages efficiently but also would learn to communicate with any group with whom he was out of contact. ERUs with low or barely average intelligence would experience considerable frustration, because their need to learn language would be greater than their ability. Many in this group would either overachieve in language (resulting in tendencies towards pedantry, malapropisms and word salads), or, defensively, would become overly taciturn and monosyllabic.

The primary factors that may be traumatic for the ERU child include (1) growing up in a socially active, gregarious group where he runs a good chance of being overlooked, ignored, slighted, and misunderstood (e.g., as an only child in a socially isolated family he may be the center of attention and the recipient of considerable understanding, while as one child in a large family he may be neglected, and thus develop negativism and hostility toward his parents and siblings); (2) strong curtailment of his activeness and expressiveness (e.g., thwarting or overcontrolling his movements). This is frequently the lot of the ERU because of the "objectionable" quality of his activity. Since, as an R, he perseverates an objectionable activity, he may be punished severely. This results in frustration and anger, usually precipitated as negativism or defensiveness; and (3) threatening or disturbing already established involvements (an emotional involvement once made is very important to him and any threat to it is personally dangerous). The ERU child is often very possessive of his mother, for example, and jealous and hostile regarding her relations

with others, including members of the family.

With this assumed dynamic base line, the second series of subtests (A, S, and PC) gives clues as to how the individual has responded to environmental pressures and as to the extent to which he has managed, controlled, or repressed aspects of his original tendencies. In general, the E child (—D) must learn to be less distractable, less responsive to immediate external events and less physically or behaviorally active. The R child (+BD) must learn to be less self-involved, more perceptive and more versatile in his responsiveness. The U child (—PA) must learn to avoid rejection, develop some type of social adaptability, and curb his negativism. In addition, as he matures, he must develop some type of facade with which to face the world; the third series of subtests (I, C, and OA) are believed to yield clues as to how he has accomplished this. Within our sample of five men, all have +I and +C, and four have +OA. (For psychodynamic reasons too complex to explain here, it is difficult to find an individual with —PC unless he is also —OA). To some extent, then, four of these five individuals should have superficially common surface personality characteristics, even though basic differences in vulnerabilities to stress can be recognized from the clues provided in the placement of the three middle tests (A, S, and PC). Finally, before a complete interpretation of any of these individuals may be made, it is necessary to have some indication of his activity level. That is, how aggressive is he? How intense is he in maintaining his adjustment? What is his threshold to discouragement and depression? The remaining test (DS) used in the PAS is believed to be the clue to these factors. Four of our five examples have been chosen with —DS while one (JC 1027) is relatively similar to JMT 52430 except for the presence of +DS.

The following chart shows the five cases in relation to one another, using JMT 52430 as a constant frame of reference:



We will use JMT 52430 as our point of departure, and note that he shows indications of having dealt with his ERU tendencies in the following ways: He has made little modification of his tendency toward responsiveness to immediate external events ($-D -A$), but has managed to control his distractability ($-D -A +I$). He has been able to retain and recall general information items comparable with his NL—thus some intellectual discipline is present. He has made some attempt to control his physical and behavioral activity through modifying his activity level ($-D -A$ with $-DS$), remaining responsive and nonaggressive. He has not modified his tendency toward self-involvement and can be expected to be somewhat lacking in perceptive sensitiveness and versatility ($+BD -S$). However, he has learned to perform perceptual and procedural tasks appropriately for his NL ($+BD -S +C$). He has learned to avoid rejection and has developed some type of social adaptability ($-PA +PC$), and is comfortable with his social adaptability ($-PA +PC +OA$). He has also learned to curb his negativism, though not to control it completely ($-PA +PC +OA$ with $-DS$). To sum this up, JMT 52430 is an experientially-oriented (practical) individual who is probably socially confident but inclined to be somewhat dogmatic, selfish, and self-centered in his interests and beliefs. Under pressure, he would be inclined to be short-tempered, with the possibility of his becoming negativistic.

JC 1027 is markedly similar in many ways. He has made little modification of his tendency towards responsiveness and has controlled his distractability ($-D -A +I$). However, his physical and behavioral control is in question. He is either an overactive, aggressive individual or a rigid, tense, overcontrolled person ($-D -A +I$ with $+DS$). Actually, he has overachieved on Digit Symbol, which is two points above his NL, and it may be expected that something intense is present here. He has not modified his tendency towards self-involvement and can be expected to be lacking in perceptive sensitiveness and versatility ($+BD -S$). In fact, he has overachieved also on Block Designs, which are two points above his NL, and it may be expected that he will be rigidly self-centered. However, he has learned to perform perceptual and procedural tasks appropriately for his NL ($+BD -S +C$). But since he has done this on a base of marked insensitivity and literalness, he can be expected to be literal or rigid in his behavior. When the $+BD -S +C$ of this intensity is combined with the $+DS$, overreactivity is ruled out and he may be expected to be markedly overcontrolled. He has learned to avoid rejection and has developed some type of social adaptability ($-PA +PC$) and is in some sense comfortable with his adaptability ($+OA$). He has a tendency to be aggressive and overactive, but overcontrols this, so it would be expected that he also has a marked tendency toward negativism or hostility ($+DS$). This also cannot be expressed because of his tense overcontrol, and will result in marked disguised hostility being present. In summary, JC 1027 is an experientially-oriented (practical) individual who is unrealistically socially confident. He is dogmatic, selfish, and self-centered in his interests and beliefs, and literal and ritualistic in maintaining his adjustment—primarily to avoid being overactive and aggressive.

JMT 54284 also has many characteristics in common with JMT 52430. He too has made little modification of his tendency

toward responsiveness to immediate external events ($-D -A$). He has also controlled his distractability and retains and recalls general information items consistent with his NL ($+I$). His physical and behavioral activity is also somewhat modified, though his aggressiveness is somewhat higher than JMT 52430 ($-DS$, but not quite so low as JMT 52430). However, he has modified his tendency toward self-involvement and has developed some capacity for versatility and perceptual sensitivity ($+BD +S$). He has also learned to perform perceptual and procedural tasks appropriately for his NL ($+C$). However, because he is $+BD$, there may be some literalness in the way he has overcome his self-centeredness and self-involvement. This usually takes the form of defined or moralistic attitudes about right and wrong, and an absence of true insight or subtleness in understanding. He has learned to avoid rejection and has developed some type of social adaptability ($-PA +PC$) and is confident about his social relations ($+OA +S$). He has also learned to curb his negativism, but can feel perfectly comfortable in being aggressive (or even hostile) when he feels it is "right" to do so ($-PC +PC +OA$ with $+S -DS$). In summary, JMT 54284 is an experientially-oriented (practical) individual who is socially confident, but inclined to be literal and precise in his need to be proper or right. He may be expected to overcontrol his temper and activity whenever he feels he is being selfish or operating in his own interests. However, he can become quite aggressive and short-tempered when he feels he is justified.

JMT 59028, on the other hand, begins to show some marked behavioral differences. He has strongly modified his tendency toward responsiveness to immediate external events. In fact, he has overachieved on Arithmetic, which is two points above his NL, suggesting that this modification has been achieved with considerable intensity. Consequently, he may be expected to be strongly defended against his reactivity and

will show many defensive features in his personality. He has also controlled his distractability and is capable of marked personal discipline when exposed to strong distraction. However, since he has not actually overachieved on the Information test, there are no indications that he has replaced his tendency toward external responsiveness with intellectual overachievement or intellectual preoccupation—a characteristic often present in the $-D +A$ individual. He has controlled his tendency toward overactivity and aggressiveness ($-D +A -DS$), and can be expected to be a somewhat passive, definitely noncommunicative individual. He has not modified his tendency toward self-involvement and has changed little of his lack of versatility and imperceptiveness ($+BD -S$). However, he has learned to perform conventional perceptual and procedural tasks appropriately when it is necessary or he is so inclined ($+BD -S +C$). He has learned to avoid or accept rejection with indifference and is comfortable with whatever social adaptation he has made, although it is more likely he will make a defensive or independent adjustment in his social-interpersonal relations ($-PA +PC +OA$ with $-DS +A$). He has learned to curb his negativism, but feels comfortable being negativistic and hostile when he is so inclined ($-PA +PC +OA$ with $-DS +A$). In summary, JMT 59028 is a tough, independent person who is highly individualistic and self-sufficient. He is defensive in his relations with others but comfortable with his social adaptability and makes only superficial efforts to modify his behavior to suit others. He follows his own interests, which are probably narrow and circumscribed. He is taciturn and insensitive, sometimes to the point of being cruel. He becomes aggressive and hostile whenever his independence and self-sufficiency are interfered with.

RAP 1017 also shows marked behavioral differences. He has not modified his tendency toward responsiveness ($-D -A$) but does have some type of control over his dis-

tractability ($-D -A +I$). He has not modified his tendencies toward self-involvement and self-centeredness and, since he shows a marked drop on Similarities, which is five points below NL, these tendencies should be quite intense. His lack of versatility and imperceptivity is also excessive. He has developed the ability to perform conventional procedural and perceptual tasks, but in a ritualistic and noninvolved way ($+BD -S +C$ with $-PC$). He will have a tendency to misperceive, and thus may perform these tasks, on occasion, inappropriately ($-PC$ with $-D -S +C$). He has not learned to avoid rejection and has not developed social adaptability ($-PA -PC$). However, he is very anxious about his interpersonal relations ($-PA -PC -OA$), and does have some ability to control his negativism; his activity level is relatively low and it is difficult for him to be aggressive or active ($-DS$ with $-PA -PC -OA$). Since he has a tendency to misperceive, and the area of his greatest anxiety is in social-interpersonal relationships, he is most likely to misperceive in these relationships. It will be expected that his social relations will tend to be distorted and interpreted by him in a highly self-centered and self-involved way ($-PA -PC$ with $+BD -S +C$). Because he is $-D -A$ he is unable to block out his responsiveness, and so he is incapable of anything but temporary schizophrenic-like withdrawals. In summary, RAP 1017 is a bewildered, anxious, and highly self-involved individual with considerable lack of insight or understanding. He is not out of touch with reality (because he is $-D -A -PC$ rather than $+D -A -PC$) but he does distort and confuse reality, particularly in his social-interpersonal relations.

Some Model Experiments

One of the more devastating studies of the potential value of pattern analysis of Wechsler profiles has been Cohen's (1955) demonstration that even skilled and sympa-

thetic clinicians could not discriminate effectively even such broad categories of diagnosis as neurosis, schizophrenia, and brain damage. This need not have been the outcome, as is demonstrated by a PAS-oriented analysis of part of the same data.⁵ Starting with a different sample of 125 cases all diagnosed as schizophrenic, the cases were classified according to their standard PAS basic patterns, and the 64 possible patterns were then ranked according to the raw frequencies of schizophrenics classified into them. Then the patterns for the 200 neurotic and schizophrenic cases of the Cohen sample were classified and assigned the ranks determined by the initial sample. Now, the top 100 Cohen cases were found to contain twice as many schizophrenics as neurotics. This result was achieved without any attempt to optimize the indicia used, and without even considering the patterns typically shown by neurotics. It seems clear that the standard Wechsler profile *does* contain enough information to permit highly valid classification into such a rough dichotomy.

This result reinforces a major implication of the discussion just above, namely, that probably the most relevant information is contained in *subpatterns* that we might extract from the total profile. It is an interesting empirical question to determine just what order of interaction must be considered in order to be certain that this information is brought under systematic scrutiny. We have begun to develop an answer to this question by carrying out contingency analyses and partitioning the degrees of freedom for individual interaction effects. In one such analysis each of 1358 cases was classified as "high" or "low" according to each of 10 profile elements derived from the WAIS and, in addition, according to 1 of 5 diagnostic rubrics. According to the null hypothesis, the distri-

⁵ Cohen's neurotic and schizophrenic cases were all tested using WB-I, but the brain-damaged were tested using WB-II. The PAS does not regard these as parallel forms, and finds WB-II inadequate for interpretive purposes.

bution of the diagnoses should be uniform in all possible profile categories. If the null hypothesis is rejected, then certain profile shapes may be associated with certain diagnoses more than with others. It has turned out that the interactions are statistically significant up to as high an order as we have been willing to go in utilizing computer time, namely, the fifth order interactions. Studies of this kind do not appear to offer an economical route to the delineation of the specific relationships in which we are interested but, with high probability of success, they may permit an unambiguous demonstration that effects of the kind discussed above are meaningful and stable.

A third approach that we have begun to explore using Wechsler data is a generalization of the moderator variable (Saunders, 1956). Beginning with the idea that the interpretation of each variable within the profile is a function of the total context provided by the rest of the profile, we can write a generalized regression equation in which all the possible interaction terms are present. By certain matrix manipulations of this equation, it is possible to define a set of "canonical moderator variables" having two desirable properties—(1) that they can be ranked in order of importance, because they are found as the Eigenvalues of a matrix, and (2) that they do not moderate one another. This simplifies the form of the equation without loss of generality. It seems likely, from the distributions of latent roots obtained so far, that the number of canonical moderators may be taken to be the same as the number of primary predictors.

Application of this technique is expected to lead to individualization of the correlation matrix of primary predictors, so that it may be estimated specifically for the region of the measurement space adjacent to a given profile. Conventional analysis of such individual correlation matrices may then lead to the determination of meaningful individual personality structures. This step has not yet been accomplished. However, we have been able to demonstrate, even in

samples of rather restricted IQ range, that the correlations of all the Wechsler subtests tend to decrease as IQ increases, and that the first canonical moderator variable appears to be directly identifiable as the NL itself. There also appears to be some stability in the configurations of weights associated with canonical moderators beyond the first, although we are not yet prepared to attempt their specific identification.

A fourth approach, that has been explored rather more extensively than the three just mentioned, is "syndrome analysis" (Saunders and Schucman, 1962). This procedure attempts to classify profiles directly into relatively homogeneous groupings, according to criteria based either on distance in the total measurement space, or distance measured in an appropriate subspace. Several sets of rather interesting results have been obtained via this approach (Schucman et al., 1962; Saunders, 1964b; unpublished data). That it has been possible to discover wider ranges of validity for some of these empirically derived groupings speaks well for both the method and the content areas to which it has been applied. However, these results also are difficult to assimilate into the existing framework of the PAS without extensive interpretation.

A modification of the syndrome approach has been developed to exploit the Wechsler collection as a whole. By specifying any possible profile and any desired subspace spanned by the measurements of intellectual performance, it is possible to obtain a ranked list of the N most similar profiles in the collection, together with identifying and descriptive information about these cases. "Unknown" profiles may be classified on the basis of such results, if due regard is paid to the base rates of the proposed diagnosis in the collection and in the population. Using "known" target profiles, the best results have been obtained when the measurement subspace is defined by NL and subtest deviations from NL, rather than by the untransformed subtest scores. We have

been dismayed to discover, however, that even 12,000 is not enough cases in the reference collection to fill the space adequately and to avoid blind spots.

It would appear that it will ultimately be necessary to superimpose a theoretically derived classification on the collection, and that the PAS may provide the basis for such a theory.

Emergent Thoughts

It would appear that both classification and measurement of individuals are worthwhile processes. Of the two, classification is the more fundamental, because the meaning of a given measurement may vary discontinuously from one class to another, while there is no such ambiguity associated with classification itself. Classification based on arbitrary dichotomizations of continuous variables does not carry this implication, of course, and perhaps ought to be known by some other name. True classification, to serve its function, must *precede* measurement.

So far as the area of intellectual functioning is concerned it is far easier to classify individuals on the basis of what the PAS calls their "ideal" type, than on the basis of either their "attained" or "primitive" type. There are both theoretical and empirical grounds for this statement. Both primitive and attained patterns consist of measurements of *status*, arrived at through the action and counteraction of many forces—some genetic and some environmental. Approximately normal distributions should and do result. On the other hand, the ideal pattern simply identifies the *direction* in which the individual is striving. So far as the elements of the pattern are concerned, this direction can only be toward or not-toward—an essentially qualitative distinction.

It would appear most fruitful, therefore, to commence a study of individual differences by classifying persons according to their patterns of interest and aspiration.

Having defined a series of such homogeneous subpopulations, we would then propose to apply a conventional factor analytic model to the study of intragroup variability, and so to discern the relevant intellectual-temperamental structures for each subpopulation in turn. We would expect to find discontinuous alterations in the nature of these structures as we move from one subpopulation to another. Actually, this strategy of research is not too different from what is commonly taking place anyway, except that we are making a virtue out of others' necessity.

Within particular subpopulations, we would expect to find that particular measures are unimportant to make, either because they exhibit no variability or because they are fully predictable from appropriate combinations of other measures, and despite the importance and potential independent contributions of these particular measures in other subpopulations. In this way, we are able to account for the overall necessity of some 18 dimensions of measurement, and yet also account for the possibility in individual cases of providing fairly complete and valid interpretations on the basis of only the subtest scores employed in the standard PAS.

REFERENCES

- BALINSKY, B. "An Analysis of the Mental Factors of Various Age Groups From Nine to Sixty." *Genet. Psychol. Monogr.* 23, 191-234, 1941.
- BIRREN, J. E. "A Factorial Analysis of the Wechsler-Bellevue Scale Given to an Elderly Population." *J. consult. Psychol.* 16, 399-405, 1952.
- CATTELL, R. B. "Personality and Motivation: Structure and Measurement." Yonkers-on-Hudson: World, 1957.
- COHEN, J. "A Factor-analytically Based Rationale for the Wechsler-Bellevue." *J. consult. Psychol.* 16, 272-277, 1952.
- COHEN, J. "The Efficacy of Diagnostic Pattern Analysis With the Wechsler-Bellevue." *J. consult. Psychol.* 19, 303-306, 1955.
- COHEN, J. "A Factor-analytically Based Rationale for the Wechsler Adult Intelligence Scale." *J. consult. Psychol.* 21, 451-457, 1957.
- COHEN, J. "The Factorial Structure of the WISC at Ages 7-6, 10-6, and 13-6." *J. consult. Psychol.* 23, 285-299, 1959.

- DAVIS, P. C. "A Factor Analysis of the Wechsler-Bellevue Scale." *Educ. psychol. Measmt.* 16, 127-146, 1956.
- GAULT, U. "Factorial Patterns of the Wechsler Intelligence Scales." *Aust. J. Psychol.* 6, 85-93, 1954.
- GITTINGER, J. W. "Personality Assessment System." New York: Human Ecology Fund, 1964.
- GOLDFARB, W. "An Investigation of Reaction Time in Older Adults." *Teach. Coll. Contr. Educ.* No. 831, 1941.
- GUILFORD, J. P. "The Structure of Intellect." *Psychol. Bull.* 53, 267-293, 1956.
- HAGEN, E. P. "A Factor Analysis of the Wechsler Intelligence Scale for Children." *Dissertation Abstr.* 12, 722-723, 1952. (Abstract.)
- HAMMER, A. G. "A Factorial Analysis of the Bellevue Tests." *Aust. J. Psychol.* 1, 108-114, 1950.
- HARONIAN, F. J. Unpublished data, 1962.
- HOVER, G. L. "An Investigation of Differences in Intellectual Factors Between Normal and Neurotic Adults." *Dissertation Abstr.* 11, 423-424, 1951.
- KODAMA, H., SHINAGAWA, F., and INDO, T. "WAIS: Japanese Edition." Tokyo: Nihon Bunka Kagakusha, 1958.
- LANFELD, E. S. and SAUNDERS, D. R. "Anxiety as 'Effect of Uncertainty': An Experiment Illuminating the OA Subtest of the WAIS." *J. clin. Psychol.* 17, 238-241, 1961.
- MCNEMAR, Q. "On WAIS Difference Scores." *J. consult. Psychol.* 21, 239-240, 1957.
- O'CONNOR, J. "Born That Way." Baltimore: Williams & Wilkins, 1928.
- RAPAPORT, D. "Diagnostic Psychological Testing." Vol. I. Chicago: Year Book Publishers, 1945.
- SAUNDERS, D. R. "Moderator Variables in Prediction." *Educ. psychol. Measmt.* 16, 209-222, 1956.
- SAUNDERS, D. R. "On the Dimensionality of the WAIS Battery for Two Groups of Normal Males." *Psychol. Repts.* 5, 529-541, 1959.
- SAUNDERS, D. R. "A Factor Analysis of the Picture Completion Items of the WAIS." *J. clin. Psychol.* 16, 146-149, 1960 (a).
- SAUNDERS, D. R. "A Factor Analysis of the Information and Arithmetic Items of the WAIS." *Psychol. Repts.* 6, 367-383, 1960 (b).
- SAUNDERS, D. R. "Further Implications of Mundy-Castle's Correlations Between EEG and Wechsler-Bellevue Variables." *J. Nat. Instit. Person. Res.* 8, 91-101, 1960 (c).
- SAUNDERS, D. R. "Digit Span and Alpha Frequency: A Cross-validation." *J. clin. Psychol.* 17, 165-167, 1961.
- SAUNDERS, D. R. "Factor Analysis of Comprehension and Similarities Items From the WAIS." Multilithed, 1962.
- SAUNDERS, D. R. "Evidence for the Relative Primitivity of Certain Traits." APA symposium on "Multitrait, Multilevel Personality Assessment—Theory, Measurement, Evaluation." Philadelphia: Sept. 2, 1963.
- SAUNDERS, D. R. "The Factorial Structure of the Kodama-WAIS for Males." (*Jap. J. educ. Psychol.*) 12, 436-438 and 508-512, 1964 (a).
- SAUNDERS, D. R. "Some Varieties of Schizophrenia." (Preliminary report.) APA symposium. Los Angeles: Sept. 4, 1964 (b).
- SAUNDERS, D. R. and SCHUCMAN, H. "Syndrome Analysis: An Efficient Procedure for Isolating Meaningful Subgroups in a Non-random Sample of a Population." Meeting of Psychonomic Society. St. Louis: Sept. 1, 1962.
- SCHUCMAN, H., SAUNDERS, D. R., and THETFORD, W. N. "An Application of Syndrome Analysis to Subjects With Ulcerative Colitis." *Amer. Psychologist* 17, 359, 1962. (Abstract.)
- SIMKIN, J. S. "An Investigation of Differences in Intellectual Factors Between Normal and Schizophrenic Adults." *Dissertation Abstr.* 11, 448-449, 1951. (Abstract.)
- SMITH, D. E. P., and CARRIGAN, P. M. "The Nature of Reading Disability." New York: Harcourt, Brace, 1959.
- STRONG, E. K. "Vocational Interests of Men and Women." Stanford: Stanford University Press, 1943.
- THURSTONE, L. L. "Multiple-factor Analysis." Chicago: University of Chicago Press, 1947.
- WECHSLER, D. "The Measurement of Adult Intelligence." (3d ed.) Baltimore: Williams & Wilkins, 1944.
- WECHSLER, D. "Manual for the Wechsler Adult Intelligence Scale." New York: Psychological Corp., 1955.

A Typology of Self-Descriptions¹

Morris I. Stein, Ph. D.² and John Neulinger, Ph. D.³

Our purpose in this paper is to describe our experiences with a typological anal-

¹ The studies covered in this paper were conducted over a period of time and at different periods were supported by a contract from the Peace Corps (PC W 54) and grants from the Carnegie Corp. and the National Institute of Mental Health (MH 10899-01) and a Public Health research career program award (MH-k6-18,679) to the senior author.

² Professor of Psychology, Research Center for Human Relations, Department of Psychology, New York University, New York, N.Y.

³ Research Assistant Professor of Psychology, Research Center for Human Relations, Department of Psychology, New York University, New York, N.Y.

ysis of subjects' responses to a self-description questionnaire based on Murray's (1938) need system. Our explorations with this technique are relatively recent and although the results are promising they need to be regarded with caution since our work is still in progress. We shall first describe the questionnaire and the types derived from it and then discuss some of our reasons for becoming interested in a typological analysis of our data.

The Self-Description Questionnaire

The Self-Description Questionnaire, developed by the senior author in a study of creativity, consists of 20 paragraphs, each of which describes one of Murray's⁴ manifest needs. The needs included are: abasement, achievement, affiliation, aggression, autonomy, blameavoidance, counteraction, defendance, deference, dominance, exhibition, harmavoidance, infavoidance, nurturance, order, play, rejection, sentience, sex, and succorance. (These 20 needs were selected because they were regarded as potential inhibitors or facilitators of creative activity.) The descriptive paragraph for need abasement, as an example, is: "I passively submit to external forces. I accept injury, blame, criticism and punishment; I surrender. I am resigned to fate. I admit my inferiorities, errors, wrong-doings or defeats. I blame myself."

In responding to the questionnaire, the subject is asked to rank the paragraphs from the one which is most descriptive of himself (rank of 1) to the one which is least descriptive of himself (rank of 20).

Both the structure of this questionnaire and the manner in which the subject is asked to respond to it distinguish this questionnaire from others. This questionnaire consists of paragraphs referring to several potential manifestations of a need and not single sentences that refer to be-

haviors or personality characteristics. It avoids the problem often encountered in similar self-description tests of tying a need to some specific form in which it might manifest itself. Thus, for example, an individual might not check an item in a personality questionnaire as being true of himself, not because he does not possess the need that the item is designed to tap, but rather because he does not manifest that need in precisely the form in which it is stated.

By using paragraphs as items, as is done in the Self-Description Questionnaire, both specific as well as general manifestations of the need in question are tapped. We draw on the individual's capacity to abstract or generalize about his own behavior or personality. When an individual has to describe himself, we assume that he does not react to himself on the basis of one single event (especially if this event is taken out of context) but scans (or has scanned) his behavior in a wide variety of situations from which he abstracts several generalizations.

We assume that an individual engages in similar behavior when he reads the paragraphs. He can extrapolate from the specific items in the paragraph the general meaning of the statements and assess the congruence between these generalizations and his assessment of himself. He can also select from these paragraphs specific statements, if these touch directly on activities in which he has engaged. While the foregoing may be an advantage of this questionnaire, it may also be a limiting factor in the range of individuals to whom it might be administered. Thus far, the questionnaire has been used primarily with individuals who are of above-average intelligence and who may be presumed to have the abstracting capacity referred to above. Where the limits are with respect to the intelligence and abstracting capacity required for the meaningful completion of this questionnaire is a matter for future research.

⁴ Murray, H. A., et al., *Explorations in Personality*, New York: Oxford University Press, 1938.

The second feature which distinguishes this Self-Description Questionnaire is the ranking procedure employed. The use of this ipsative method is dictated by the theoretical maxim that an individual's personality is best reflected not in the intensity of his separate motives as compared to those of others but in the *organization* of these motives.

The Types

Method

The Self-Description Questionnaire was administered to three different populations and identical statistical procedures were used in the analysis of each group's data.

The three groups were:

1. *Peace Corps*.—The subjects were 80 male volunteers accepted for training and further selection with the Peace Corps in an overseas program of community development. The men in this group were all in their midtwenties, and all but one had some college education; about half had completed college.

2. *Creativity*.—The subjects were 116 male Ph.D. chemists, employed in an industrial research organization, who participated in a study of creativity (Stein, unpublished). Most of the subjects were in their midthirties. They came from three different industrial organizations and consisted of: 31 "more creative" subjects, 34 "less creative" subjects, and 51 subjects who formed a "middle group." Creativity evaluations were based on ratings by superiors, peers, and subordinates.

3. *New York University (NYU) Students*.—The subjects were 115 male students at New York University enrolled in introductory Government classes. Most of the subjects were between 18 and 21 years old.

The Self-Description Questionnaire was administered to all subjects in group sessions. In spite of the fact that it is not a simple matter to rank 20 paragraphs, most subjects completed the questionnaire in 20 to 30 minutes.

Results

The subjects' rankings of the 20 paragraphs were the data used for Q-technique factor analyses.⁵ Eighty by 80 (Peace Corps), 116 by 116 (Creativity), and 115 by 115 (NYU) matrices of product-moment coefficients of correlation between subjects' need rankings were computed for Q analyses. Principle components of each of the intercorrelation matrices were extracted and rotated by means of the Varimax method.

The Q-factor analyses of the data obtained from the different populations yielded the following results: four factors for the Peace Corps population (types I, II, III, and IV), five factors for the Creativity population (types A, B, C, D, and E) and seven factors for the NYU population (types N, O, P, Q, R, S, and T). Differences between the means for each of the 20 needs on each of the aforementioned factors within each of the three populations were tested by analyses of variance. Only two needs out of 20 in each of the populations were found not to discriminate significantly among the types. For each of the populations these needs were: Peace Corps—harmavoidance and counteraction; Creativity—order and counteraction; NYU—abasement and infavoidance. Accordingly, all 20 needs were included in the final type profiles.

⁵ Twenty by 20 matrices of product-moment coefficients of correlation between individual paragraph ranks were computed for R analyses. Principle components of each of the intercorrelation matrices were extracted and rotated by means of the Varimax method. The factors obtained indicate that, within each population (Peace Corps, Creativity, and NYU), some needs tend to cluster. Certain patterns of need covariations emerge but they are not consistent across populations. For example, a factor composed of "sentience-play-sex" (positive) and "defendance" (negative) emerges from the analysis of the Creativity population data; however, these needs load maximally on different factors when the data for the Peace Corps are analyzed. In the latter case, sex and play load low on one factor, sentience on another, and defendance on a third. Similarly, in the NYU population, the four needs show loadings on four different factors.

In view of the above, nothing would be gained by using factor scores instead of the original need scores for any further applications of the test. The data indicate that the inclusion of correlated needs in the test introduces relatively little redundancy.

To establish type profiles, attention was focused on "type-definers." These were subjects who loaded highest on any given Q factor and substantially higher on this factor than on any of the other factors. Data from these type-definers were used to establish both the profile of the type and its limits. The profile of a type is determined by the type-definers' mean rankings of the needs. The distribution of these rankings is used to set the limits for the type. A two standard deviation range around the mean of each need was selected as a permissible area of fluctuation. In this manner the type definition takes into account the relative importance of each of 20 needs as a discriminator for any given type. A particular need may have considerable fluctuation in rank in one type, whereas for another this same need may be limited to a relatively small range.⁶ For example, the limits on counteraction for type C are ranks 1 to 7, based on a mean of 4.2 and a standard deviation of 1.8; the limits on counteraction for type E are ranks 1 to 12, based on a mean of 4.3 and a standard deviation of 3.9. Thus, the rank which a person gives to counteraction is more critical for type C than it is for type E. In the extreme case, given a large enough standard deviation, the rank a person assigns to a need may become irrelevant for that type.

To "type" new individuals, two approaches have been developed. One is a graphic technique in which a profile is made up of a new subject's ranking of the needs. This profile is then matched to transparent overlays of the type profiles and type limits. Considering the number of types that have been identified, this involves 16 matchings for each individual. A more rapid and economical procedure involves the use of a computer. The data

for each of the 16 types have been programmed, and a rapid typing of new individuals is possible. In both methods, the criterion for a match is the degree to which a person falls within the limits of the type. If he falls completely within these limits, he is categorized as belonging to that type. He also belongs to the same type if he exceeds these limits on one or two needs.⁷ Exceeding the limits on more than two needs eliminates him from that type category. Should a subject satisfy these conditions on two or more types, he is categorized as belonging to that type on which he exceeds the limits the fewest number of times. And, in the case of ties, use is made of the subject's deviation scores from the profiles of the types. Categorization is then determined by the lowest sum of *z* scores over the 20 needs.

Obviously, not all individuals can be categorized according to available types. Hopefully, the number of such unclassified individuals will be decreased as new type profiles are developed.

Combining Types

Some of the 16 types obtained in the three separate Q analyses may be considered sufficiently similar to be grouped together. At least two different approaches can be utilized in deciding whether to establish such combinations. One is statistical, and the other is psychological. In the former, some measure of quantitative similarity is used, while the latter is determined by the psychological similarity between the types based on the interpretative significance of the need hierarchy, as well as other psychological test data and behavioral information that might be available for the individuals in the types. It is conceivable that the profiles of two types may be highly correlated, but psychodynamically they might be rather different. Under such circumstances, it

⁶ Present values of type means and standard deviations are subject to sampling and measurement error and will obviously need to be adjusted in the future. However, the principle of a type definition in terms of mean and standard deviation will likely be maintained.

⁷ The broadening of the type category was undertaken to achieve greater similarity between typing through Q-technique factor analysis and the computer technique.

would be wise to keep the profiles separate and distinct.

Currently we are pursuing both statistical and psychological approaches to effect combinations. Product-moment correlation coefficients between type profiles (table 1) and "unadjusted coefficients of correlations" of the same data show, basically, the same clusters of types with but minor differences. To illustrate how the psychological approach is used in conjunction with the statistical approach, consider the combination III-N-A-D. Following the statistical data, these types cluster; however, psychological analysis suggests that it would probably be wise to maintain III-D as one cluster and A-N as a separate cluster. The psychodynamic analyses of these types based on the rank order of the needs within the types as well as some knowledge of test scores and observations of individuals within the types suggested that type AN was more self-directed and more likely to be open to both internal and external experience than is the combination IIID. This difference, considered from a psychological point of view,

was deemed a critical enough basis for separating the AN and IIID types.

Additional support for the type combinations comes from experience in typing new individuals. If one determines a person's second best type fit, one usually arrives at one of the other types in the appropriate type combination. And when type-definers of the Creativity and Peace Corps populations were inserted as "markers" in the Q analysis of the NYU population, they fell into types, the similarity of which was suggested by the correlational procedures described above.

We have concentrated in most of our studies on the five types that have showed up with greatest frequency in the populations we have studied.⁸ Rather reluctantly and aware of the danger of such labeling, we have named these type combinations as follows: Socially Oriented (IBR), Intellectually Oriented (IIEQ), Action Oriented (IIID), Unconventional (IVP), and Resourceful

⁸ In fact there are 6 type combinations (IBR, IIEQ, IIID, IVP, AN, and CO), two specific types (S and T), and one "Unclassified" (X) category.

TABLE 1.—Product-Moment Correlations (Pearson r) Between 16 Type Profiles

	II	E	Q	^a III	N	A	D	I	R	B	IV	P	C	O	S	T
II		.82	.89	.63	.66	.74	.36	.58	.56	.29	.57	.63	.39	.45	.33	.23
E			.84	.65	.65	.64	.63	.73	.68	.48	.45	.49	.39	.31	.24	.32
Q				.46	.56	.54	.29	.59	.51	.31	.45	.48	.29	.31	.41	.40
III					.79	.81	.78	.58	.58	.16	.32	.39	.53	.64	.21	.33
N						.86	.61	.76	.58	.39	.55	.46	.32	.33	.45	.13
A							.63	.53	.48	.30	.47	.49	.48	.53	.43	.11
D								.59	.57	.47	.07	.09	.46	.35	.03	.22
I									.88	.66	.44	.33	.02	-.05	.44	.45
R										.58	.29	.30	.18	.13	.36	.38
B											.09	-.05	.00	-.34	.08	.50
IV												.85	-.18	-.01	.42	.08
P													.07	.26	.40	.18
C														.86	-.10	.22
O															.06	.02
S																.30

^a Bracketed correlations are suggested type combinations.

(AN). The order of the needs in the various types is presented in table 2.

Brief psychodynamic descriptions of the types follow:

Type IBR: Socially Oriented.—The individual in this type appears to be dedicated to other people. He enjoys being with and working with others. In doing so, he finds greatest satisfaction in assisting helpless individuals and in supporting, comforting, and protecting others. A man of this type is also sympathetic. He avoids hurting others and will not be critical or severe in his interpersonal relationships. He is not a hostile or aggressive person. Compared to the other individuals we studied, a person of this type is less autonomous and more deferent to authority. He can be dominated by others and is likely to accept restrictions placed on his behavior. Indeed, he will strive to conform to others' wishes; when he knows what they want, he will try to fulfill their requests.

Type IIEQ: Intellectually Oriented.—Members of this type also enjoy being with other people and working with them. Unlike the members of type IBR, however, they place greater emphasis on achievement. In striving to attain their goals, these men are

likely to follow internal frames of reference. They believe they know what is best. Therefore, if they and the authority figures they work for are not in agreement, conflict may develop because they want to follow their own inclinations. If there is agreement with their superiors these volunteers can be quite loyal and devoted subordinates. They are sensitive and aware of both internal needs and external stimuli. Compared to the other four types, they are more likely to seek out and enjoy sensuous impressions and to enjoy esthetic feelings. But since they are not very well organized, they may find themselves frustrated in utilizing their esthetic impressions constructively. Left to their own devices, members of this type are likely to leave a situation which is not to their liking, rather than cope with the difficulties they encounter, and seek a new environment which will give them greater satisfaction.

Type IIID: Action Oriented.—An individual of this type aims to get things accomplished. He enjoys working with others, but it is most important to him to be in control of the situation. Compared to the other four types, an individual of this type is more involved with achievements and with dominating social situations. A man of

TABLE 2.—Order of Needs in Different Types

Socially Oriented	Intellectually Oriented	Action Oriented	Unconventional	Resourceful
Affiliation	Achievement	Achievement	Sex	Achievement
Counteraction	Affiliation	Affiliation	Affiliation	Affiliation
Achievement	Autonomy	Counteraction	Autonomy	Counteraction
Deference	Counteraction	Dominance	Play	Play
Nurturance	Sentience	Nurturance	Nurturance	Order
Order	Nurturance	Order	Succorance	Dominance
Sex	Sex	Deference	Sentience	Sex
Blameavoidance	Order	Exhibition	Exhibition	Sentience
Succorance	Succorance	Autonomy	Counteraction	Exhibition
Harmavoidance	Rejection	Defendance	Defendance	Nurturance
Play	Play	Play	Order	Autonomy
Infavoidance	Dominance	Aggression	Dominance	Deference
Defendance	Deference	Sentience	Deference	Harmavoidance
Sentience	Blameavoidance	Sex	Achievement	Blameavoidance
Autonomy	Defendance	Succorance	Infavoidance	Defendance
Dominance	Infavoidance	Harmavoidance	Aggression	Succorance
Exhibition	Harmavoidance	Rejection	Abasement	Infavoidance
Rejection	Aggression	Infavoidance	Blameavoidance	Aggression
Abasement	Abasement	Blameavoidance	Rejection	Rejection
Aggression	Exhibition	Abasement	Harmavoidance	Abasement

NOTE—The needs that are bracketed are tied for the same position.

this type sees himself as confident and as liking to influence others but, at the same time, he is aware of others' needs and can be quite nurturant and sympathetic. He is inclined to be systematic in what he does and accept responsibility. He is not likely to avoid situations because he might be blamed for his actions. A man of this type gives higher priority to fulfilling goals than to satisfying his impulse life. He is a doer and an achiever rather than a player. When he does satisfy his impulses, he will do so in a socially acceptable and approved manner, rarely overindulging himself, for he would not like to appear inferior. He gives the impression that he will strive to be upright and sincere in whatever he does by following a code that he has been taught.

Type IVP: Unconventional.—A person of this group enjoys going to the aid of helpless people with whom he can be supportive and comforting and from whom he apparently receives support and comfort in return. A man of this type is not very discriminating in his interpersonal relationships and he is unlikely to reject others unless they try to dominate him. A member of this group likes to have others around him so that he can be seen and heard, can entertain and amuse. He enjoys playing and indulges his impulses. He will seek out others who will provide him with sympathetic understanding and possibly some direction and leadership, but not domination, for this man prides himself in not being abusive or submissive to others. He is not very well organized, for he sees organization as coercive and does not see order as allowing for freedom. Nor is he an achieving person, for he regards himself as a free soul who would be just as happy to see others go their own way too. Underneath it all he may be thankful that others are more organized, because it is through their presence that he can go off and satisfy his own impulses. Finally, if others are kindly disposed toward him and provide him with leadership in a permissive atmosphere, he

will, under these conditions, be able to develop and achieve goals.

Type AN: Resourceful.—A person of this type is oriented to achievement and is not passive. He enjoys being with others but is not likely to be submissive to them. Indeed, in a situation where there is a choice between satisfying his own or others' needs, he will elect to satisfy his own first. He is capable of being easygoing and relaxed and enjoys relationships with women. He feels under pressure when his freedom is interfered with or if he feels that his position of mastery is threatened. Under such circumstances he feels anxious, but he soon recovers his resources and finds new means of coping with his problems. Thus, a person of this type is rather flexible and willing to accept momentary compromises as he keeps his eyes on long-range goals.

Reliability

The reliability of the Self-Description Questionnaire has been studied in two ways: (a) rank order correlations for the rankings of the needs, and (b) a coefficient of agreement, kappa.⁹ Kappa, in this instance, involves the study of the frequency with which members of a type continue being members of that type on subsequent test occasions. Kappa represents the percentage of agreement corrected for chance. Test retest reliability data in terms of rank order correlations were obtained from entering college freshmen who were tested over a 4- to 6-week period at the beginning of their first semester. For males ($N=80$) an average correlation of 0.74 and for females ($N=68$) an average correlation of 0.79 was obtained.¹⁰

The data for kappa using the same population of subjects with nine possible cate-

⁹ Cohen, J. "A coefficient of agreement for nominal scales." *Educ. Psychol. Measmt.*, 1960, 20, 37-46.

¹⁰ The results were obtained as follows: Rank order correlations were computed separately for each subject. These rhos were then converted into z' equivalents, averaged, and then converted back into rhos. With the 20 paragraphs in the Self-Description Questionnaire taken as $N=20$ observations, a rho of 0.53 is significant at the 0.01 level.

gories (six type combinations, two specific types and one unclassified category) were 0.35 for males (N=80) and 0.25 for females (N=68). Both kappas were significant beyond the 0.001 level.

Frequencies of the Types

To date our experience with the types has been limited primarily to young people in the United States, most of whom are attending college. The relative frequencies with which the different type combinations appear in these groups—one group at New York University and the other a Peace Corps population—are indicated in the last two columns of table 3. (The other data in table 3 will be discussed later.) From this table it is apparent that four type combinations account for most of the subjects. These combinations are: Socially Oriented, Intellectually Oriented, Unconventional, and Resourceful. It should also be noted that we cannot classify all of the individuals who answer the Self-Description Questionnaire. We were unable to type 19 percent of the students in an entering class at NYU and 5 percent of a large population of Peace Corps trainees.

The fact that we cannot classify all individuals is a problem that will concern us in future research. Obviously, our typology system can be extended to include more types. But such types may turn out to be of lesser and lesser frequency in “normal” populations, and it becomes a matter of

utility how specific we want to become in our classification system. On the other hand, the very aspect of atypicality may well be related to nonconformity, as data reported on in the next section suggests.

Studies Employing the Types

The types have been employed in a series of investigations designed to explore their validity as well as their usefulness in understanding behavior.

Experimental Studies

Neulinger¹¹ recently investigated the behavior of Socially Oriented and Intellectually Oriented types in a study related to dissonance. One purpose of the study was to demonstrate the need for a classification of subjects prior to any experimental manipulation. It was shown that different types of individuals react quite differently to the same environment, i.e., experimental manipulation. The Socially Oriented subjects generally expressed greater negative affect with regard to a task that they had to perform. This negative affect was manifest in greater expression of fatigue and difficulty ratings of task. A possible explanation of their behavior was their greater “evaluation apprehension” (Rosen-

¹¹ Neulinger, J. Person types, environment types and resultant forces. New York University: unpublished doctoral dissertation, 1965.

TABLE 3.—Percent of Different Populations in Various Types

Type name	Type code	Japan (N=74)	Poland (N=95)	Italy (N=200)	Israel (N=106)	New York University (N=412)	Peace Corps (N=1133)
Socially Oriented.....	IBR	38	22	16	17	21	28
Intellectually Oriented.	IIEQ	12	9	12	26	19	20
Action Oriented.....	IIID	4	1	4	5	2	3
Unconventional.....	IVP	5	11	21	14	18	12
Resourceful.....	AN	1	3	2	14	10	28
	CO	4	9	4	6	7	2
	S	0	3	2	2	3	2
	T	0	7	2	4	2	1
Unclassified.....	X	34	34	36	12	19	5

berg)¹²—that is, their desire to achieve a more positive evaluation from the experimenter. They also behaved in a more conforming manner as manifest in the degree to which they agreed with certain statements. Neulinger also found that being called upon to perform an unpleasant task for which one has not volunteered is not particularly inconsistent with a Socially Oriented self-image, while for an Intellectually Oriented person this state of affairs can be quite disturbing (dissonance arousing). This results from the fact that the Intellectually Oriented person is higher on need autonomy than is the Socially Oriented person. From what has just been said one might surmise that Socially Oriented subjects might volunteer with greater frequency for psychological experiments. In fact, however, this is not the case. Among 47 Socially Oriented subjects who were asked to volunteer for Neulinger's experiment, 22 (47 percent) refused but of 30 Intellectually Oriented subjects only four (13 percent) refused. The discrepancy in frequency in volunteering may be accounted for by the greater apprehensiveness of the Socially Oriented subject and the greater intellectual curiosity of the Intellectually Oriented subject. This finding underscores the fact that researchers might be well advised to study the types of individuals they are working with lest they overgeneralize from their results. This might be particularly true of research in academic settings where the investigator often has to work with student volunteers as subjects. Results, with such subjects, rather than being relevant as general principles of human behavior, may in fact be limited to certain types of subjects.

We have recently had an experience which supports this latter suggestion. Frances Berger (unpublished) in our laboratory replicated an experiment by Pepi-

tone¹³ in which he found that subjects in whom high self-evaluation was induced were more competitive than subjects in whom low self-evaluation was induced. While the results for the total population did not corroborate Pepitone's findings, when we look at the different types we find that the Unconventional type does follow Pepitone's hypothesis. They are more competitive under high self-evaluation conditions than under low self-evaluation conditions. The results for the Socially Oriented type run counter to the hypothesis. They become more competitive under low than high self-evaluation conditions. Although this latter result is inconsistent with Pepitone's general thesis, it is consistent with a finding in another of Pepitone's¹⁴ studies where he suggests that low self-evaluation subjects who see a connection between the self-evaluation condition and the competitive game situation may use the game as a means of changing the experimenter's evaluation of themselves and therefore behave more competitively. It may be that the Socially Oriented type regards low self-evaluation as a greater threat to its security system than does the Unconventional type, and that they therefore react more competitively under these conditions than when they are highly evaluated.

Neulinger's finding regarding the conformity of the Socially Oriented type was recently corroborated in two experiments by Ellman (unpublished) in our laboratory. The first was a replication of the autokinetic experiment.¹⁵ In this experiment Type B's (as a member of the Socially Oriented type) estimates of the apparent movement of a light were significantly more effected by a confederate's estimates than

¹³ Pepitone, A. D. *Attraction and hostility: An experimental analysis of interpersonal and self-evaluation*. New York: Atherton Press, 1964.

¹⁴ Pepitone, *vide supra*.

¹⁵ Sherif, M. A study of some social factors in perception. *Arch. Psychol.*, 1935, 27, No. 187.

Linton, Harriet. Autokinetic judgments as a measure of influence. *J. abnorm. soc. Psychol.*, 1954, 49, 464-66.

¹² Rosenberg, M. J. When dissonance fails: On eliminating evaluation apprehension from attitude measurement. *J. Pers. soc. Psychol.*, 1965, 1, 28-42.

were the estimates of other types studied. In the second study a variation of the Asch¹⁶ experiment type B's were similarly significantly more influenced by confederates' judgments than were other types. An interesting aspect of both experiments was that those subjects whom we were unable to classify were among the least conforming. This suggests the possibility that a common denominator to both their relative lack of conformity and nonclassifiability of their Self-Description patterns may represent a kind of atypicality that is worthy of further investigation.

Types in a Cross-Cultural Study¹⁷

We have also explored the types in a cross-cultural study. Our purpose was to study the relative frequency with which the types would appear in foreign countries. The countries studied were Japan, Poland, Italy, and Israel, and the populations studied were professional or university groups. The obtained data (table 3) indicate several items of interest.

First, either the types we have found and have currently been working with are not sufficient to account for all kinds of persons or they are culture bound. In all likelihood both of these are probably accurate. The insufficiency of the types will in all likelihood be rectified as our research proceeds but the evidence for the fact that they may be culture bound is reflected in the fact that we were unable to type one-third

of three of the foreign populations—Japan, Poland, and Italy. We were unable to classify only 12 percent of the Israeli population. If we assume that Israel is more westernized or more Americanized than the other populations, then this can account for the greater classification of Israelis.

The possibility of characterizing nations through the relative frequency distribution of the types (i.e., "modal personality patterns"¹⁸) is suggested by table 3. The most frequent type in the Japanese group was the Socially Oriented type and it appears in this group with a greater proportion than in other populations. This is to be expected since the Socially Oriented type is the most deferent of our types. In the Polish population the modal group is also the Socially Oriented type but here it accounts for only about one-fifth of the population. Among the Italians the most frequent type is the Unconventional type. This type is most acceptant of impulses and again agrees with the stereotype one might have of the Italians. Finally, one might also have predicted the most frequent type among the Israelis to be the Intellectually Oriented type. Obviously we do not present these data to suggest that we have done a study of national character, but merely to suggest that the Self-Description Questionnaire could be utilized in such a study.

Assessment Studies

In addition to the laboratory and cross-cultural studies of the types, the types have also been utilized in two major assessment studies, one of creativity and the other of the first group of Peace Corps volunteers in Colombia, South America.

Types and Creativity

One hundred and sixteen Ph. D. industrial research chemists were subjects of our

¹⁶ Asch, S. E. Studies of independence and conformity: I. A minority of one against a unanimous majority. *Psychol. Monogr.*, 1956, 70, No. 9 (Whole No. 416).

Blake, R. R. and Brehm, J. W. The use of tape recording to stimulate a group atmosphere. *J. abnorm. soc. Psychol.*, 1954, 49, 311-313.

¹⁷ We wish to thank the following individuals who collected the data for this study: Miss Yachiyo Kato and her associates at the Society of Science and Man, Tokyo, Japan; Prof. Gabriele Calvi, Istituto di Psicologia, Università Cattolica del Sacro Cuore, Milan, Italy; Prof. Mieczyslaw Choynowski, Head of Department, Polish Academy of Sciences, Psychometrical Laboratory, Warsaw, Poland; and Prof. Joel Shanani, Department of Psychiatry, Hadassah Hospital, and Department of Psychology, Hebrew University, Jerusalem, Israel.

¹⁸ Inkeles, A., and Levinson, D. National character: The study of modal personality and sociocultural systems. In Lindzey, G. (ed.) *Handbook of Social Psychology*. Cambridge, Mass.: Addison-Wesley Publishing Co., 1954.

creativity research. They made up three groups in terms of the ratings on creativity they received from their superiors, peers, and subordinates. One was a group of "more creative" men ($N=31$), the second was a group of "less creative" men ($N=34$) and the third was a middle group ($N=51$). Of the 65 chemists in the more and less creative categories we found 23 more and 25 less creative individuals who were members of our types. Both more and less creative men appeared in all types. Of the 23 more creative men, 48 percent are of type A, 7 percent of type B, 34 percent are equally divided between types C and D, and 7 percent are of type E. Of the less creative men, 16 percent are of type A, 36 percent of type B, 12 percent of type C, 24 percent of type D, and 12 percent of type E. As indicated by these data, some types yield more creative individuals than do others. The largest proportion of more creative men appear in type A, and the largest proportion of less creative men appear in type B ($X^2=7.84$, $p<0.02$). In all the other types combined there are almost equal proportions of more and less creative individuals.

The study of the types in relation to creativity also allowed us to explore the relationship between the types and some biographical information. There is much biographical information that we have on the types which will be analyzed in the future. At present, we should like to present data on only one aspect of parent-child interaction for the two types A and B.

In attempting to understand how more creative individuals develop, we investigated the possibility that they were exposed to complexity early in life. One source of early complexity that we thought would exist was the extent to which the subject perceived his mother as inconsistent in her relationship with him. It was assumed that an inconsistent mother might be frustrating to the child, and the child, to structure his own environment or to satisfy his needs, would be thrown more onto his own re-

sources than a child reared in a consistent environment. And having had this experience of using his own resources would stand him in good stead in future creative work.

To gather the necessary data, a questionnaire entitled "Interpersonal Relations in Childhood"¹⁹ was utilized. In this questionnaire subjects were asked to rate, on a seven-point scale, the degree of consistency or inconsistency they recall having perceived in their mother. The item read: "As a child I felt my mother was:" and then the rating was to be indicated on a continuum that ranged from very inconsistent to very consistent.

It was found that the more creative chemists ($N=31$) did indeed regard their mothers as less consistent ($H=5.04$, $p<0.05$) than did the less creative men ($N=34$). Consequently, if inconsistency of the mother was related to creativity status, it should also differentiate between type A and type B groups. Here we find a trend in the direction of the hypothesis. The type A group does tend to rate their mothers as less inconsistent than does the type B group ($H=2.65$, $p<.10$).

Types and the Peace Corps²⁰

The second assessment study in which the types were utilized was the assessment of the first group of Peace Corps volunteers who were assigned to a community development program in Colombia, South America. One of the problems on which the type approach helped cast light was the difficulty the assessment group had in predicting the behavior of certain individuals. In this situation 80 young men had taken the Self-Description Questionnaire and 73 (91 percent) could be classified according to the five major types. The data indicated that all individuals in the Action Oriented type were accepted for service in Colombia but

¹⁹ Developed by Prof. Joel Shanan and the senior author.

²⁰ The Peace Corps study is described at greater length in Stein, M. I. *Volunteers for Peace*, New York: Wiley, 1966.

that among the other types between 20 to 35 percent were rejected. Using acceptance/rejection rates as an indication of the assessment group's preferences, it is apparent that it lay with the Action Oriented group. The assessment group also made predictions as to how effective the different individuals would be. These predictions, however, did not differentiate between the types. After the volunteers completed their 2 years of service in Colombia, ratings on their effectiveness were obtained from their superiors. Compared to all of the other types, the type that had the greatest probability of acceptance, the Action Oriented type, received the lowest of the effectiveness ratings.

As we study the psychodynamic descriptions of the types in an effort to understand why the Final Selection Board may have fallen short in their predictions we can develop a series of speculations. We would speculate that the Final Selection Board overevaluated the self-confident air of the Action Oriented volunteers, and that they underestimated the other types in the following ways: There may have been some doubt that the Resourceful type could work in a team endeavor, that the Unconventional type would be sufficiently well organized, that the Intellectually Oriented volunteer might not have his intellectual interests satisfied, and that the Socially Oriented volunteer might have been too deferent and too lacking in autonomy. Whether these were the specific reasons for errors in the assessment prediction we cannot say but the speculations just presented alert us to possible difficulties that other assessment groups may have with the same types of individuals in other situations.

The relationships between types and roles were also investigated in the Peace Corps study. The volunteers were part of a community development program. Basically, this program concerned itself with involving Colombian villagers in the organization of community councils called *juntas* which then had the task of trying to fulfill various community needs. In

participating in community development activities volunteers could assume several different roles such as the role of a Technician, an Intellectual Idealist, or a Community Developer, etc. The roles that the volunteers had assumed were classified by the volunteers' superiors. Several case studies of how the types fulfilled the Community Developer role suggest that although individuals of different types fulfill the same role they vary in their role behavior or style. Here are several instances of responses of several different types of volunteers to the same problem—that of organizing the *junta* or community council.

An Action Oriented volunteer said:

"I sent out invitations to all the *campesinos* (peasants) in the area. I advertised the *junta* (council) organizing meeting. I told the schoolchildren to bring their parents to the meeting. Then I contacted all the influential people in the area and explained to them what community development was, then had them push the *campesinos* into turning up at the meeting. Thus to get the people out we did it by written invitation, their children told them of it, and pressure from their neighbors."

A Socially Oriented volunteer said:

"For a number of weeks beforehand I visited the individuals and families in the area in order to have closer contact with prospective *junta* members. When I was satisfied that there was enough interest in organizing a *junta*, I sent out mimeographed notices to all members of the area (through schoolchildren) besides having the *padre* (priest) notify them and request their attendance."

An Intellectually Oriented volunteer reported the following:

"In one instance, we visited each house in the area during a several-day period, trying to explain what we were up to. We also had the *padre* announce the proposed meeting at Mass. On the day of the meeting we walked through the village, and the people followed us to the schoolhouse. The *promotor* (Peace Corps volunteer's

Colombian counterpart) gave a long explanation about community development, and then the *junta* was named (no real election) with everyone in agreement."

These three instances reflect the relationships between personality type and style of behavior. All three of the volunteers were regarded as Community Developers yet they differed in that the Action Oriented volunteer emphasized forcefulness and power, the Socially Oriented volunteer emphasized contact with people and their readiness, while the Intellectually Oriented volunteer emphasized explanation.

Summary

This paper has concerned itself with a number of personality types based on Self-Description data. We have discussed ways in which the types have been studied to investigate the validity of their psychodynamic descriptions and several instances of their usefulness in clarifying data obtained in both laboratory and assessment studies. Because of the small number of subjects involved in these studies the results have to be regarded as tentative rather than definitive.

There still remain several important problems. Assuming the nonrandom distribution of ranked self-descriptions, i.e., assuming that such a distribution represents something resembling a multimodal, multidimensional distribution, is Q-factor analysis the appropriate or the optimal technique in identifying these modal peaks? We are about to investigate the statistical appropriateness of our techniques. Whatever the outcome and consequent changes in our classificatory system, further explorations will continue along the path described in this paper. Additional issues to be investigated are the biographies of the types, their approaches to learning, problem solving, their capacity for change, and the like.

In concluding this paper we should like to make more subjective comments based

on our experiences as they relate to the broader concerns of this conference.

We believe that a typological system can be of real value to the understanding and prediction of behavior. Psychology has long emulated the physical sciences by studying the relationships between isolated variables. This approach has obviously yielded much significant data and information. In addition, however, we also need studies in which variables are related to each other within the organization of the individual's personality. The variable to variable approach implies that an individual's motivation is reflected in the intensity of his separate motives. Personality theory, however, suggests that his motivation is best reflected in their organization. Two individuals, for example, may have precisely the same score on a measure of need achievement. For one, achievement may have an aggressive and forceful quality, while for another it lacks any compulsive quality and is indeed associated with playfulness and pleasure. These may become important differences depending on the conditions under which studies are conducted.

Secondly, the typological approach is more consistent with another principle of personality theory—equifinality. Different kinds of individuals can achieve the same end result by following different pathways or processes to the goal. For example, consider the body of research in the area of creativity. Much of this research includes some statement of the personality variables that differentiate individuals who differ in creativity status. From this research, it is easy to gain the impression that creative individuals are all of one sort and less or noncreative individuals are of another. A typological analysis, as the one indicated in this paper, reveals that different types of individuals can occur in both groups. In all likelihood what differentiates them are their styles of work and the appropriateness of a given style for any specific situation.

The typological analysis of individuals seems of potential value in another area. By calling attention to the possibility that there are types of individuals, one is alerted to the fact that many of our theories of personality that strive to be theories of man are probably in actuality theories of men. This is nothing new,²¹ but by calling attention to it through the concept of types we are better able to determine areas of overlap and areas of uniqueness in different theories.

As we have become more involved in our work with typologies we have become more aware of several general problems. Obviously classification is the cornerstone of any science. No science has been able to make significant progress without a classification system. There are problems, however, in the development of such a system—some are inherent in the problem of classification while others are extrascientific. One of the most difficult of the latter to overcome is the bias that people have against classification systems. There is a feeling that the use of a classification system does injustice to the individual being classified because it overlooks his individuality. There is also the fear of the self-fulfilling prophecy. Once an individual has been classified, persons react to the classified individual in terms of the classification rather than in terms of his individuality. Both of these are real problems but one should bear in mind that these problems are not inherent in the problem of classification but in the ways in which classifications are used. Whatever system of classification is used it

has to be cross-validated in a variety of situations.

One last point that may be more critical than the previous two is that no classification system of general usefulness is likely to be developed until there is agreement among behavioral scientists as to the variables to be included in the development of the classification scheme. Until we have agreement on the basic variables or parameters, we are apt to develop systems that would be of limited utility.

Discussion

ARDIE LUBIN, Ph. D., *Research Psychologist, U.S. Navy Medical Neuropsychiatric Research Unit, San Diego, Chairman*

(Papers: "Biometric Assessment of Mental Patients," Joseph Zubin, Ph. D., Chief of Psychiatric Research (Biometrics), New York State Department of Mental Hygiene, New York, N.Y. "Patterns of Intellectual Functioning and Their Implications for the Dynamics of Behavior." David R. Saunders, Ph.D., Professor of Psychology, University of Colorado, Boulder, and John W. Gittinger, M.S., Psychological Assessment Associates, Washington, D.C. "A Typology of Self-Descriptions," Morris J. Stein, Ph. D., Professor of Psychology, Research Center for Human Relations, New York University, and John Neulinger, Ph. D., Research Assistant Professor of Psychology, Research Center for Human Relations, New York University.)

²¹ Maskin, M., has called attention to this matter as follows ("Adaptations of psychoanalytic technique to specific disorders," In Masserman, J. H. (ed.) *Science and Psychoanalysis*, 3, *Psychoanalysis and Human Values*, New York: Grune & Stratton, 1960, pp. 321-352.): "Freud used hysteria as the model for his therapeutic method, depression as the basis for his later conjectures. Adler's clinical demonstrations are rivalrous, ineffective, immature character types. Jung's examples were restricted to a weary, worldly, successful, middle-aged group. Rank focused upon the conflicted, frustrated, rebellious artist aspirant. Fromm's model is the man in a white collar searching for his individuality. And Sullivan's example of choice is the young catatonic schizophrenic."

FORMAL DISCUSSION OF PAPERS

	Page
ROBERT E. HARRIS, Ph. D.	404
EDWARD W. FORGY, Ph. D.	410

OPEN DISCUSSION

Drs. ZUBIN, NEULINGER, SHAKOW, and STEIN....	415
--	-----

Discussant's Remarks

ROBERT E. HARRIS, Ph. D., *Professor of Medical Psychology, School of Medicine, University of California, San Francisco.*

First some comments on the papers by Stein and Neulinger, Saunders and Gittinger, and Zubin. I have also taken the liberty of commenting on the paper by Mattsson and Gerard although it was not assigned to me. It is an early report of a major study of schizophrenia in which psychological tests were utilized, along with measurements in other domains.

Stein.—Stein offers a unique departure from the usual self-report inventories by requiring the subject to rank order 20 Murray needs for their self-descriptive salience, each need being represented by several fairly concrete statements of ordinary behavior and attitude. The rank orderings of the 20 needs form the basis for a matrix of intercorrelations among persons from which can be extracted types. These types Stein labels as Socially Oriented, Intellectually Oriented, Action Oriented, Unconventional, and Resourceful. He goes on to demonstrate that these different types are associated with different kinds of responses to experimental psychological procedures, and that they occur with different frequencies in different occupational and nationality groups. Although the method in its present form is limited to rather high educational levels because of the complexity of the task posed for the subject, it seems likely that modifications could extend its range downward into at least the upper half of the distribution of intelligence in the general population and be useful in describing many neurotic and psychotic patients. It would be interesting to see the relations of these types to other nontest criteria and to tests such as the Strong Vocational Interest Blank, the MMPI, the Edwards Personal Preference Schedule (which is also based on Murray's needs), and eventually to clinical data.

Saunders and Gittinger.—This represents an attempt in the tradition of Wechsler and Rapaport to utilize the level and patterning of Wechsler-Bellevue subtests for the diagnosis of noncognitive personality functions. In view of inconclusive findings in previous efforts with this method, and in view of McNemar's persuasive cautions on the low reliability of subtest scores and of differences between these scores, I feel that the burden of proof at this time rests on Saunders and Gittinger to show that their Personality Assessment System has important correlates in other test and nontest behavior. Their system has a certain inherent attractiveness about it, independent of the statistical methods by which it was derived. A jagged profile of scores, like an errant life history, catches the clinician's eye, and clinicians are often reluctant to accept the notion of chance. However, the assigning of noncognitive meaning to each of the Wechsler-Bellevue subtest scores seems hazardous enough, and inferring personality structure, dynamics, and history from their complex interactions as described in the case examples even more so. Almost any one of the interpretations offered would seem to require a major program of validation against external criteria, a tremendous task. Some notion of the scope of the task can be obtained by examining recent papers by Gough (1965a, 1965b) on the validation of the Socialization Scale, only one of several scales in the California Psychological Inventory. Meaning has accrued to scores on this scale by application to many different kinds of groups known to be different on one or another criterion, by correlations with other tests, clinical judgments, etc.

Zubin.—The techniques which Zubin lists under his heading of culture-dependent fall naturally into the category of rating scale methods rather than psychological testing. As culture-fair he classifies methods for studying various aspects of language behavior and its modification by reinforcement. The methods for assessing the com-

municability of schizophrenic speech seem particularly promising. Other procedures for studying language which might be mentioned are Hargreaves' and Starkweather's (1963, 1964, 1965) analyses of the speech of psychiatric patients into its noncontent, vocal, formal characteristics. These have included measures of rate and duration of utterances, and most recently they have found that the spectrum of frequencies in speech sounds shows a close relationship with clinical change in depression. Analyses of gesture and posture by Ekman (1965) also show promising relations with clinical change and clinical diagnosis, modes of expression of affect, etc.

Although it may strain Dr. Zubin's classification of language analysis as culture-fair, I would like to suggest that analysis of the *content* of speech may yield to computer methods. If computers can aid linguistic analysis to assign proper authorship to books of the Bible and to unmask ghost-writers, perhaps they will be sensitive to schizophrenic peculiarities. Attempts in this direction have been made by Starkweather and by Iker and Harway, using frequencies of words and of clusters of words. Free-running discourse in an interview may be unwieldy, but procedures like interpreting proverbs, which place more constraints on the total word sample, and on pairs and trios of words and their associative frequencies, may be more amenable. Normals probably respond with a limited number of different words and phrases to the proverb: "A rolling stone gathers no moss." Some of Benjamin's schizophrenic and organic verbalizations and other abnormalities may be readily identifiable: self-reference, concretism, incomplete and faulty desymbolization, literalness, the use of rare words, obsessive overideation, and perhaps even twisted, overly elaborate, and idiosyncratic syntax.

I am not sure what is the common denominator among the procedures which Zubin calls culture-free. Elsewhere Zubin has stated that responses which occur within

1,000 milliseconds after a stimulus may not be influenced by cultural factors. He also states that physiological, sensory, perceptual, and psychomotor responses are less likely to depend upon culture than are conceptual measures. Would he include perceptual and cognitive styles around which a substantial body of research is building? How about perceptual constancy, autokinetic movement, size estimation, embedded figures, and field articulation? Many of these procedures yield differences between schizophrenics and others. However, I see several limitations:

1. Although significant differences by a "t" test may be found, the magnitude of the relationships may be quite low as measured by correlation.

2. Findings, often based on small samples in atypical situations, may lack generality.

3. Relationships may be confounded by the process-reactive, paranoid-nonparanoid, and acute-chronic dimensions.

4. Relationships among the several procedures tend to be low, thus providing poor definitions of the underlying dimensions.

The relation of deviant performance in these laboratory situations to the major phenomena of schizophrenia, for example, depersonalization and delusions, is obscure. I would like very much to hear Dr. Zubin and also Dr. Shakow discuss possible connections between these procedures and clinical phenomena. This might lead to a clarification of the underlying dimensions, and to the design of direct methods for measuring them.

Mattsson and Gerard.—The Michigan schizophrenia project employed psychological, physiological, and biochemical tests and clinical observations (Gerard, 1964). From a pool of more than 140 psychological tests 28 were selected, yielding an unspecified number of variables. Presumably the basis for selection was the power of a test to differentiate schizophrenics from nonschizophrenics. The authors classify the 28 tests as intelligence and thinking, memory, perceptual efficiency, and psychomotor per-

formance. It should be noted that no personality tests of the self-report or projective variety survived, or perhaps were not included in the original battery. A factor analysis reduced the battery to eight factors. It will be of great interest to examine the complete tables of intercorrelations and factor loadings in the full report of the study, because there tends to be a great deal of methods variance in such procedures, i.e., variables derived from one test procedure tend to factor out together and to be unrelated factorially to other variables from another procedure which should by name and surface appearance be related. Such factors tend to be highly specific to the procedures from which the variables with high loadings were derived, in which case they are likely not to be related to any outside criterion. If several different procedures are involved, with careful attention to eliminating highly correlated and, hence, redundant variables, then the factor loadings tend to be small and to give an inadequate definition of the dimension. The latter problem confronted Halstead (1947) when he attempted to identify his factors of biological intelligence.

Nevertheless, the psychological tests used appear to be very powerful in separating schizophrenics from nonschizophrenics and, within the schizophrenic group, the paranoids from the nonparanoids. They also contribute largely to the statistical derivation of types. If I interpret the tables correctly, it would seem that psychological tests contribute more to the typologies than do the biochemical or physiological measures. Furthermore, when the authors elaborate their descriptions of the types by referring back to all of the variables in the study (no fewer than 388!), psychological tests contribute importantly to their interpretations.

Now some more general comments. Except for the Stein procedure none of the methods described above are designed to explore directly the subjective world of the patient and his emotional, motivational,

attitudinal characteristics, the content of his thinking, his social attitudes and behavior, his familial and job relations, his habitual modes of coping with stress, etc. This is the stuff of personality study in most outpatients and newly admitted inpatients except those who are stuporous, educationally very limited, or otherwise inaccessible. Two kinds of psychological procedures are designed to explore such matters. I refer to self-report inventories and projective tests. Since they are not covered elsewhere in this conference, I shall discuss them briefly by an example of each, with an eye to what dependable dimensions they offer for personality study.

Self-Report Inventories

I shall use the MMPI as an example of self-report methods. The problem of how many and what dimensions exist within the MMPI has been a recurrent problem. Fortunately it is not necessary to rely on the dimensions which were built into the inventory according to how the diagnostic pie was sliced in one neuropsychiatric institute in the late 1930's and early 1940's. Factor analysis, profile analysis, testing of non-clinical populations, and other methods have enabled the test to go well beyond the original scales. Soon after its publication more than 20 years ago a series of factor analyses demonstrated that two factors could account for most of the common factor variance in the published scales. Some investigators concluded that the test measured nothing more than two factors: the first was called anxiety, or dysphoria or general maladjustment or something of the kind; the second was called repression or extroversion or somatization, and was a less clear factor than the first. Later factor analyses, for example by Fisher (1964), employing different methods of factor extraction and rotation, have revealed as many as six factors, four of which held up across

different kinds of patients and different clustering methods. The conclusion that the MMPI measures nothing but two factors was premature. The claim was based on particular factorial methods and furthermore, referred only to the common factor variance. There is also the possibility, because items were selected from the responses of clinically distinct groups, that each scale contains reliable specific variance.

Moreover, early in the history of the MMPI, it became clear that the published scales were themselves factorially complex. For example, within the hysteria scale one set of items describes physical symptoms and complaints, while another affirms a kind of blandness about psychologically troubling problems. Within normal samples, these two sets are negatively correlated; high scores on both sets suggest the *belle indifférence* of the classical hysteric, e.g., equanimity in the face of a paralysis of both legs. Thus the hysteria scale and other published MMPI scales are already second-order factors, and intercorrelations among them yield very general third-order factors like general maladjustment, or the likelihood of psychiatric referral for any of a number of reasons.

This logic led to an examination of the items within each of the MMPI scales in an effort to identify sets of items which have some common underlying psychological meaning. Twenty-eight such sets were identified and assigned names on the basis of their content. Subscales with enough items for adequate reliability were constructed (Harris and Lingo, 1955), and they have been found useful in interpreting MMPI profiles. The intercorrelations of these subscales, even within a single MMPI scale, range from about -0.50 to more than $+0.80$. A principal components analysis (Lingo, 1960) with varimax rotations revealed no fewer than seven factors which were replicated across four groups: male and female patients and male and female

nonpatients. An additional four factors were replicated across three groups and three more across two. Nine of these factors seemed to have sufficient generality across groups to warrant the development of new scales. These new scales are currently in the process of clinical validation.

With the development of large-capacity computers it became possible to factor analyze the intercorrelations of the individual items within each MMPI scale. Comrey (1958) has shown that there are from 5 to 16 factors within each scale. Some of these are sufficiently similar to those derived by the subjective methods described above to reassure one that the human mind is a fairly efficient data processor. In Comrey's analyses probably too many factors were retained for rotation and the factors are too small for reliable measurement. However, they further demonstrate the factorial heterogeneity within the published MMPI scales. Most recently, Tryon (1965), by newly developed methods of cluster analysis, has analyzed the entire pool of MMPI items. He eliminated 317 items which had low communalities, and further refinement reduced the pool to 292 from which he derived 7 clusters. These are: (1) Introversion; (2) somatic symptoms; (3) suspicion, mistrust; (4) depression, apathy; (5) resentment, aggression; (6) autism, dissociative thought; and (7) tension, worry, fears. Seventeen-item scales to measure these factors show reliabilities ranging from 0.81 to 0.91, and slightly longer scales, reliabilities of 0.85 to 0.94.

Thus the MMPI contains many dimensions derived from different methods and levels of analysis. These dimensions may well be important in exploring a wide range of psychopathology and, one way or another, in classifying people.

Projective Techniques

Turning now to projective techniques, it should be said at the outset that these

methods have not been as productive of solid research findings as their wide clinical use would lead one to expect. In a review of projectives, Gleser (1963) discussed some of the problems in validating these methods and reported many negative results as well as a few intriguing positive ones. Certainly we are a long way from demonstrating the validity of projective testers' rich interpretations of schizophrenic, neurotic, and other kinds of psychopathology. It seems likely that we have not yet discerned the important dimensions within a patient's responses to ambiguous stimuli; we have few ready-made, scaled dimensions to utilize in a study such as the Michigan schizophrenia study. Observations of the behavior of clinicians as well as a review of the literature suggests that the conventional scoring categories may not be the most incisive ones nor the ones on which the psychologist bases his interpretations.

Two recent studies by Dr. Margaret T. Singer illustrate meaningful dimensions within projective materials. In a study of the Rorschachs of patients with ulcer and ulcerative colitis, Krasner and Kornreich (1954) found that among the many Rorschach categories only the total number of responses differentiated the groups. This was consistent with the slightly higher IQ of the colitis patients. In the face of these negative results Dr. Singer undertook a blind sorting of the protocols on the basis of her previous extensive experience with ulcer and other psychosomatic groups. She was correct in 50 out of 54 of the cases. Other people utilizing the cues she identified as the basis for her judgments were able to achieve better than chance discriminations. The underlying dimension is difficult to define, but involves viewing the total patient-ink-blot-tester transaction as a sample of the patient's communication processes.

Second, in a series of studies with Wynne (1963, 1965) of the families of schizophrenic

patients, Dr. Singer has developed scoring criteria and a manual for a dimension of thinking disorder which ranges from amorphous through fragmented to stably constricted. Scoring of the Rorschachs of the parents of schizophrenics, neurotics, and normals permitted a prediction of the clinical diagnosis of the schizophrenic patient and inferences about the form of his thinking disorder. Dr. Singer's approach to projective test materials is illustrated in the following:

As lists of mechanically totalled Rorschach scores, or as TAT interpretations preoccupied with content of stories, this negative valuation of projective techniques for family studies would be emphatically correct. However, in this study, projective techniques have been used quite differently. The test protocols were viewed as a sample transaction between subject and tester, not simply as projections of intrapsychic problems, not primarily as a glimpse at the subject's intrapersonal frame of meanings through which he or she interprets reality. The transaction between subject and tester provides a relatively standard way of sampling attention, thinking-communication and relating. Thus, the protocols can be used as a means of studying the same kinds of stylistic aspects of thinking as have been observed in the clinical work with the families. Indeed, we have used the same categories, characterizing patterns of dealing with attention and meaning, for rating both projective test protocols and family therapy excerpts, two distinctly different forms of data.

Regarded as sample transactions, rather than as tests in a narrow sense of the word, the projective techniques thus provide a convenient but not a unique source of data which has several advantages when properly obtained: (a) the contribution of the investigator to the transaction is relatively standardized, both in the cards used and in the kind of communication made by a well-trained examiner; (b) the subject is free to interpret and structure the

stimuli and the test situation with a wide range of possibilities, not foreclosed by the nature of the procedure itself; (c) experienced psychologists have a vast knowledge of the norms and variations in projective test transactions to draw upon in the interpretation of specific findings, whereas less is known about expected performance in other freely running transactions.

These methods for scoring Rorschach behavior are being extended to family interactions in a social situation, using the same scoring categories. It is important to note that the dimensions that Dr. Singer has isolated are defined in conventional psychiatric language, or at least in a language which can be common to: (a) clinical psychiatric description, (b) descriptions of social interactions, and (c) psychological tests, especially the projectives which allow considerable scope for behavior, verbal and otherwise, in a semistandardized situation. It may be expected that with common dimensions and a common language, links between different methods of personality study will be more easily verifiable.

To close: This long digression into self-report and projective tests is meant to suggest simply that these methods complement measures of cognitive, perceptual, and motor functions, and conventional methods of psychiatric examination. They contain meaningful dimensions along which individuals may be arrayed in explorations of the total psychopathological space.

REFERENCES

- (1) COMREY, A. L., and MARGGRAFF, W. M. "A Factor Analysis of Items in the MMPI Schizophrenia Scale," *Educational and Psychological Measurement*, 1958, 18, 301-311.
- (2) EKMAN, P., in TOMKINS, S. S., and IZARD, C. E. (editors). "Affect, Cognition and Personality," pp. 390-442. New York: Springer, 1965.
- (3) FISHER, J. "Some MMPI Dimensions of Physical and Psychological Illness," *Journal Clinical Psychology*, 1964, 20, 369-375.
- (4) GERARD, R. W. "The Nosology of Schizophrenia: A Cooperative Study," *Behavioral Science*, 1964, 9, 311-333.
- (5) GLENER, G. G. "Projective Methodologies," *Annual Review of Psychology*, 1963, 18, 391-422.
- (6) GOUGH, H. G. "Conceptual Analysis of Psychological Test Scores and Other Diagnostic Variables," *Journal of Abnormal Psychology*, 1965, 70, 294-302.
- (7) GOUGH, H. G. "Crosscultural Validation of a Measure of Asocial Behavior," *Psychological Reports*, 1965, 17, 379-387.
- (8) HALSTEAD, W. C. "Brain and Intelligence: A Quantitative Study of the Frontal Lobes," Chicago, University of Chicago Press, 1947, p. 206.
- (9) HARGREAVES, W. A., and STARKWEATHER, J. A. "Recognition of Speaker Identity," *Language and Speech*, 1963, 6, 63-67.
- (10) HARGREAVES, W. A., STARKWEATHER, J. A., and BLACKER, K. H. "Voice Quality in Depression," *Journal Abnormal Psychology*, 1965, 70, 218-220.
- (11) HARGREAVES, W. A., and STARKWEATHER, J. A. "Voice Quality Changes in Depression," *Language and Speech*, 1964, 7, 84-88.
- (12) HARRIS, R. E., and LINGOES, J. C. "Subscales for the MMPI: An Aid to Profile Interpretation." Unpublished manuscript, the Langley Porter Neuropsychiatric Institute, San Francisco, Calif., 1955.
- (13) KRASNER, L., and KORNREICH, M. "Psychosomatic Illness and Projective Tests," *Journal Projective Techniques*, 1954, 18, 353-367.
- (14) LINGOES, J. C. "MMPI Factors of the Harris and Wiener Subscales," *Journal of Consulting Psychology*, 1960, 24, 74-83.
- (15) SINGER, M. T., and WYNNE, L. C. "Thought Disorder and Family Relations of Schizophrenics. III. Methodology Using Projective Techniques," *Archives of General Psychiatry*, 1965, 12, 187-200.
- (16) SINGER, M. T., and WYNNE, L. C. "Thought Disorder and Family Relations of Schizophrenics. IV. Results and Implications," *Archives of General Psychiatry*, 1965, 12, 201-212.
- (17) TRYON, R. C. "Unrestricted Cluster and Factor Analysis, With Applications to the MMPI and Holzinger-Harman Problems." Unpublished, obtainable from author, University of California, Berkeley.
- (18) WYNNE, L. C., and SINGER, M. T. "Thought Disorder and Family Relations of Schizophrenics. I. A Research Strategy," *Archives of General Psychiatry*, 1963, 9, 191-198.
- (19) WYNNE, L. C., and SINGER, M. T. "Thought Disorder and Family Relations of Schizophrenics. II. A Classification of Forms of Thinking," *Archives of General Psychiatry*, 1963, 9, 199-206.

Discussant's Remarks

EDWARD W. FORGY, Ph. D., *Research Specialist, Statistics, Department of Psychiatry, Medical Center, University of California, Los Angeles*.¹

I would like to organize my remarks on the papers presented here around a response to the basic point made earlier by Dr. Goldine Gleser and later by Drs. Grinker and Nunnally in almost the same words. Paraphrased slightly, it is:

Typologies are unnecessary when the measurement data itself is available. Grouping into classes always loses information. Types therefore lead to poorer predictions and decisions than if the same data were related directly to the relevant criteria.

Since all three papers in this session have either developed or look forward to developing typologies that are explicit functions of measurement data, they seem to lie directly in the line of fire. Therefore it seems pertinent to reconsider reasons for the possible value of such work.

There is, I believe, a potential gain from such activities. However, the possible benefits seem unlikely to be realized unless investigators are quite explicit about the purpose or function of the systems they develop. It has been persuasively argued that drastically different functions cannot ordinarily be accomplished adequately by the same classification system. It would follow that, if systems are developed by analysis of empirical data, then the nature of the grouping procedure used should depend upon the purpose of the typology—as well as the form of the data. It will be shown that each of several major purposes calls for a different, characteristic way of handling the measurement data—even

after questions on variables, scaling, and similarity functions have been resolved.

At least three logically distinct general purposes for typologies could be stated:

Purpose I

A typology can be useful simply for convenience in communication. This modest purpose is not at all in conflict with Dr. Gleser's point. However, even though some information in a multivariate profile is lost, the convenience may outweigh the loss, particularly when the investigator or clinician needs to communicate about patients without any visual or mechanical aids. For example, among Minnesota-trained people the MMPI two-point codes² seem to serve this purpose quite well.

In general, to develop such a typology empirically, a statistical grouping procedure should minimize the inevitable loss of information by making groups as homogeneous as possible over the set of measures. Naturally, the more groups the less information is lost, but the less convenient the system is to use. Precise minimization of any loss function (for a given number of groups) in large samples of multivariate data is completely impractical even with computers, but some quite acceptable approximate solutions are available (e.g., Ward (1)). Actually, almost all available clustering methods try to make groups homogeneous, although it's usually not clear just what loss function is to be minimized by the final system.

But is it this purpose—mere convenience of communication—that really motivates researchers? I'm sure it's not; they expect their typologies to do something more. The next two purposes, I trust, come closer to these expectations.

Purpose II

A typology can reflect a fact of nature, that there are actually discrete, separated

¹ This paper was prepared with the partial support of the National Institute of General Medical Sciences (grant GM-11055), and of the UCLA Alcoholism Research Clinic (California Department of Public Health Contract No. 110). Computing assistance was obtained from the Health Sciences Computing Facility, UCLA, sponsored by NIH grant FR-3.

² This makes up an explicit but an intuitively defined typology. It uses a semiordinal transform of the t-scores, in terms of which the types are perfectly homogeneous.

subtypes of individuals within the larger sample. It is widely recognized that even when data are essentially continuous measures with marginal distributions that appear ordinary and unimodal, the joint distribution can have distinct multiple modes—that is, be “clustered.” When this occurs it may be an extremely important discovery. A good example is provided by the astronomers Hertzsprung and Russel, who observed a striking case of multimodality after plotting stars on the two dimensions of temperature and luminosity (5). This discovery was repeatedly verified, and has strongly affected theories of stellar evolution. A description in terms of types—if the types correspond to separated clusters—can lead to an improved understanding of nature.

On the other hand, the description of nonclustered samples in terms of discrete types will probably add only to the substantial confusion already in our field. If we are to avoid proliferation of pseudo-discoveries, the methods used to detect natural typing in data must also be able to tell us, “No” when there is none. In the course of trying to work on this extremely difficult problem myself, I carried out some Monte Carlo experiments with several popular clustering procedures upon artificial samples of nonclustered data points (Forgy (2)). Appalling to say, most of the methods would always say, “Yes,” and chop up a sample from a single joint normal population into some number of “types.” How often has this happened to others without their realizing it?

So, for this purpose, typological methods can be two-edged instruments; at this point in time I rather suspect that the wrong edge is getting used most of the time!

Purpose III

A typology can enhance our ability to relate to or predict an outside criterion. Actually, as Dr. Gleser pointed out, this would be impossible if one were free to

explore all possible ways of using the scores directly. But the set of all possible predictive functions is infinite, and many require enormous sample sizes to be used at all,³ so this is not really a practical alternative to typology. In practice, though, the familiar set of linear functions is nicely limited, and its use is certainly easy with today’s computers and libraries of programs. Before we become very excited about the relation of a particular typology to a particular criterion, I think we should at least compare its performance to that of multiple linear regression (or multiple discriminant analysis, if the criterion is nominal) using the same measures. This will show whether the typology is telling us more—or less—than the measures do directly.

If relations are strongly nonlinear it’s possible indeed for the typology to come out ahead, and such a result could be of both practical and theoretical interest. But is it likely? We’ll never know until it’s done—hopefully soon, and hopefully to some of the data presented at this conference. One reason why the likelihood may be fairly low is that the clustering methods so far used have not been directed toward any explicit outside criterion; no *Y* variables have entered into the grouping process, so it would be surprising if the resulting classification were very efficiently related to them. What seems to be needed, to produce typologies that are at all likely to come out ahead of linear predictions, is

- (a) remarkably intuitive a priori definitions of types, or
- (b) methods to construct typologies empirically which maximally relate to specific criteria.

Just as maximally homogeneous types can’t be solved for exactly, but can approxi-

³ Direct estimation of multidimensional outcome surfaces, without any restrictive assumptions, is particularly costly. A modest experiment using 10 scores, each expressed merely as high-medium-low, would require cases in each of $3^{10}=59,049$ cells for each treatment group.

mately, computing algorithms can probably be developed to locate a partition into almost maximally relevant types.⁴ When found, their degree of relevance (e.g. proportion of *Y* variance accounted for) can be computed and compared to that from linear prediction, both for data already used and for cross-validation samples. A found superiority of typology might be short-lived, because inspection of it might suggest an appropriate transform to cure the nonlinearity, but it would have served a useful purpose of improving prediction and perhaps also understanding of the phenomenon.

Another use of typology—within purpose III, but calling for yet other clustering procedures—could be to enhance, rather than compete with, prediction via linear methods. An explicitly defined set of classes within which linear prediction is optimal could be much more effective than either process by itself. Such classes might or might not correspond to any separated clusters evident in the predictor data. While the concept of piecewise linear regression has been proposed before (Hanken (3)), methods for implementing it leave much room for improvement.

It is obvious that a typology quite relevant to one criterion may be irrelevant to another. Purpose III is therefore really a broad class of purposes which share a common methodological problem. This problem is actually the general one of nonlinear prediction, and the use of typologies constitutes one approach to it.

To summarize: Three general purposes of typologies are distinguished, each calling for a different kind of statistical grouping method. Typologies from the different approaches would respectively

- (I) have maximally homogeneous classes,
- (II) correspond to naturally discrete groups,

(III) be maximally relevant to an outside criterion.

Some choice of objectives is necessary, because in most data these three will conflict with each other. Specifically, there may be no discrete groups at all, but relatively homogeneous classes are always possible. Also, the division into separated clusters will not necessarily maximize homogeneity—as it would not in the Hertsprung-Russel data, for instance (5). Or, there may be naturally discrete groups (e.g. male-female) but ones which are irrelevant to a particular outside criterion. Or, most of the variables (producing most of the variance) may be utterly irrelevant to the outside criterion, so maximum homogeneity and maximum validity cannot be achieved by similar partitions.

Most available clustering methods appear to be roughly suitable for the first purpose, which unfortunately is the least important scientifically. They are very probably misleading for the second purpose, and again very probably quite ineffective for the third. Although several dozen grouping procedures have been proposed, there is still an acute need for methods that give results really pertaining to the interesting questions of discrete subtypes and improved relevance or prediction. In the meantime, the investigator with realistic expectations from statistical grouping methods will probably contribute more useful information to his field than one without.

Remarks on Specific Papers

Dr. Zubin's paper is chiefly concerned with the preliminary tasks of developing a more adequate set of relatively objective and quantitative variables, but he definitely looks toward using these measures to obtain a classification system for patients that is an overall improvement upon the standard psychiatric groupings. The principle aim, besides that of assigning patients to classes

⁴ For the concept of maximally relevant groups, I am indebted to Dr. James B. MacQueen of the Western Management Sciences Institute, UCLA.

more objectively and therefore more reliably, appears to correspond to purpose I above. The most frequent criticism of the traditional diagnostic system was that there is far too much heterogeneity within diagnostic categories, so the new system would presumably emphasize homogeneity within its own groups.

Although it stated, "If new classification categories emerge, it will be a gain for science," I am sure it's better classification that's desired, not just an alternative—which could even be worse. Therefore, one might reasonably ask: How can the relative goodness (in terms of homogeneity, say) of the new classes versus the old be compared? The new classes will certainly be more homogeneous with respect to the set of measures used in their own development, for any clustering method will make a *post hoc* partition that cannot help but take advantage of errors of measurement, etc. And—paradoxically—the less related are the new measures to the old classifications, the greater will be the apparent advantage of the new system over the old! It is obvious that other criteria would be necessary for a fair and objective comparison of old versus new, but what would they be?

Perhaps homogeneity is not all that is really desired. Earlier the search for discontinuities in patients' behaviors or characteristics is also seen as rewarding, and this corresponds to our purpose II—discovering which natural clusters, if any, exist. It might be very rewarding indeed, but it will call for a very different way of handling the data.

Purpose III (establishing groups relevant to some outside criterion) is perhaps not left out altogether either, since among the faults of the standard diagnostic system are low relationships to course of illness and outcome. Again, a different grouping process would be needed to maximize the relevance of classes to these outside criteria, and this classification system would almost certainly differ greatly from that which

emphasizes homogeneity with respect to the predictor measures themselves.

In any case, a more explicit conception of the purpose(s) of the classification system(s) would seem to be necessary before reasonable choices can be made on the next steps of the project.

The development of reliable indices and scales of pathology stands as a worthy contribution in its own right. One would hope that their use is not limited to classification studies, because it may be some years before "the new techniques for fractionating populations into homogeneous subgroups" can confidently be relied upon to make very efficient—or even reasonable—use of the data given them.

Dr. Zubin later remarked that some of the other results presented at this conference do give evidence that typologies can do better than the dimensional approach. I would have to disagree, and say that we unfortunately have yet to see a study that really compares the different approaches on the same data. Such studies are exactly what is needed if the new clustering techniques are to contribute anything substantial to our knowledge. What we do have now in the empirical studies is evidence that computer-derived typologies have greater than zero relevance to a number of things—but so do other kinds of computer-derived functions as well as clinical judgments and diagnoses.

Stein and Neulinger develop an interesting typology for normals and then present data on its relevance to several kinds of outside facts. The descriptive base for their classification is a narrow but logically coherent set of measures in the same domain. While such narrowness of the base might seem a weakness to some, their study helps illustrate its advantages; types here tend to be understandable, and can be described more concisely as to what they have in common. If the typological approach has advantages, perhaps it is more likely to be through the practice of building a number

of fairly modest subsystems, each based upon a different subset of measures.

The alternative is to include a wide variety of measurements (and probably larger numbers of measures) into the analysis at one time, and then to develop a single all-inclusive typology. Such groups will necessarily be less homogeneous with respect to most variables and types will differ from each other in many ways—so that found relevancies are explainable by a large number of possible hypotheses.

The psychiatric diagnostic system itself is of the all-inclusive and nested type, with several broad classes each split into scores of mutually exclusive categories. An alternative structure (which has already been suggested here by Dr. Elmer Gardner in conference paper No. 3) is a crossed system which consists of several logically independent classification variables (each of which may or may not be ordered) to describe the patient. The crossed type of system usually seems easier to work with in research and statistical studies, because crossed classifications can be reasonably collapsed to obtain larger groups for different special purposes.

Now back to the Stein and Neulinger paper. Although the test-retest reliability of the typology is rather low, it is refreshingly unusual to see any figure at all for this important parameter. The data presented are interesting psychologically, and the types are indeed relevant to nationality, creativity of chemists, and selection and performance in the Peace Corps.

However, one can't help wondering how much of the relation should be credited to the variables themselves, and how much to the typology. Are the relations because of the typology, or in spite of it? Would the variables, taken singly or in linear combination, have clearer and stronger relations to nationality, creativity, etc.? I believe answers to such questions would contribute to our understanding of classification methods and perhaps also to the author's understanding of the self-description questionnaire variables themselves. It may seem

uncharitable to wish that a researcher had analyzed the data differently when he has already come up with interesting relationships, but in the context of this conference the issue of whether a given typology adds or loses information is quite pertinent. Unless typologies add something, I think that in the long run we will be better without them. There are now at least as many clustering methods as participants in this conference, and many of these (transposed factor analysis is one such) contain within themselves a wide variety of alternatives which will shape the results (e.g. number of factors taken out, method of rotation). So it is possible for the same set of measures upon the same people to give us hundreds of different classification systems. These would give different results, both as to the degree and the kind of relations they bear to outside criteria. Unless there are interactions which can demonstrably be detected and exploited, the sole ingredient added by typologies may well be confusion.

So, I would again respectfully urge the researchers in this session to also analyze their measurement data by the usual criterion-oriented methods, to give themselves and the rest of us some perspective to the value of typologies from any source, deductive or empirical.

The summary by Stein and Neulinger contains a phrase which may explain why this sort of thing was not done: "Assuming the nonrandom distribution of ranked self-description, i.e., assuming that such a distribution represents something resembling a multimodal multidimensional distribution * * *."

Apparently the writers have made just this assumption in their work, and so are most concerned with finding the best method to describe these multiple modes. But what evidence is there for this assumption? Q factor analyses almost always seem to suggest to their interpreters that the data contains subtypes. Most other clustering methods do too, but how do we know what is really there? Multiple modes are indeed

fascinating and important, but their existence cannot be assumed from results of the kind shown here.

It would be nice if there were a dependable technique to recommend. If only two or three modes are suspected, then plotting the data points on pairs of the first several principal axes would be likely to show them. If all clusters in the data are fairly round, then a technique I developed (Forgy (4)) can discriminate between clustered and nonclustered data. But in general we need something much better than these.

In concluding their discussions, Stein and Neulinger suggest that a generally useful classification system must follow agreement as to what variables should go into it. One can only hope that this is not true, for psychiatrists and psychologists seem as likely to disagree on the variables as on the classifications themselves. Since it is unreasonable to expect any one classification system (or any set of measures) to be optimal over a wide range of uses, work on subsystems built on limited data bases and pointed toward particular purposes will probably be much more useful in the long run than trying to achieve a system that will do all things for all men. One might almost say that the less we aspire to do in classification, the more we are likely to accomplish.

REFERENCES

- (1) WARD, J. H. Jr., "Hierarchical Grouping To Optimize an Objective Function." *J. Amer. Statist. Assoc.*, 58, 236-244, 1963.
- (2) FORGY, E. W., "Evaluation of Several Methods for Detecting Sample Mixtures From Different n -Dimensional Populations," American Psychological Association Convention Paper, 1964, Los Angeles.
- (3) HANKEN, ALBERT F., "Method and Model for Analysis and Description of Car-Following Performance," Ph. D. Dissertation, Dept. of Industrial Engineering, Ohio State University, 1964.
- (4) FORGY, E. W., "Detecting 'Natural' Clusters of Individuals," *Western Psychological Assoc.*, Santa Monica, April 1963.
- (5) STRUVE, OTTO and ZELBERGS, VELTA. *Astronomy of the 20th Century*. Fig. 129, p. 259. Macmillan, New York, 1962.

OPEN DISCUSSION

(Following this formal discussion of papers by Drs. Harris and Forgy, Dr. Lubin opened the meeting for discussion from the floor and served as moderator.)

Dr. ZUBIN. Dr. Harris is still fighting the battle to save psychological tests, but I've given up. In fact that is the story of my life. In dealing with intelligence tests, when they were still in bloom, I tried to find out whether they could be used to show patients with general paresis improving under fever therapy. After that was over there was nothing left to do for intelligence tests. Then came the personality inventories. They, too, fizzled for me although some workers are still holding on to them. Then came sorting tests, and they fizzled out. And then came the Rorschach. I came to the conclusion that the Rorschach is nothing else but an interview. If that is all there is to it, why not go into the interview directly? That is why we got Dr. Burdock and Dr. Spitzer to work with the interview itself.

I am not at all surprised that intelligence tests and other tests haven't lived up to their expectations for psychopathology. As a matter of fact, some 5 years ago, since I myself was a little bit too close to the field, I got a student, Dr. Charles Windle, to review the prognostic value of psychological tests. And, as you know, that turned out to be a complete flop. There wasn't a single test that held up prognosticwise in the entire field. So that is why I have turned away from tests and turned my attention to culture-dependent, culture-fair, and culture-free indicators. Dr. Harris asks a lot of an indicator. All we are saying is that performance on these techniques are indicators, not causes of mental disorder. Do you ask, "Why does a Wasserman test work?" Does anybody know why it works? We use it because it is an indicator. However, if you want to know whether there is any way of explaining the usefulness of say, our reaction time studies, I believe there is a para-

graph in my paper which will do that. The patient seems to be overprepared by his last trials or his last experience, and that may be the reason why he behaves differently.

Dr. NEULINGER. I would like to discuss Dr. Forgy's comment. In terms of some of our findings when we thought we had established discrete groups, he says, "Are these relations because of the typology or in spite of it?" I think the answer is actually given in part by his own comments. He mentioned three purposes of a typology. One of these purposes is to establish maximally discrete groups. Dr. Forgy says he tried various techniques, and it seemed that inevitably, he came up with discrete groups as mere artifacts. That may be so. But because he gets discrete groups as an artifact is no argument that, in fact, discrete groups could not exist. We make the assumption that they do exist, and possibly methods haven't been developed yet to get at them.

If our purpose is to predict (the second purpose of a typology), it is possible for any given criterion to find some linear relationship of dimensions which would better predict for that particular criterion, than a typology which has not been developed with this criterion in mind.

Which brings me back to the issue discussed here; our typology takes into account the environment, and is based on a conceptualization which allows us to relate any criterion to the typology. Our method allows us, in fact, to type not only the person but the environment, or the criterion, in the same conceptual system (i.e., in terms of press) and then relate them to each other. Hopefully in this way we might arrive at a better prediction.

Dr. SHAKOW. I want to say something about the point Dr. Zubin made with regard to tests. I am not so pessimistic about tests as he is, especially when I see them used properly. I think it is possible to use tests such as Rorschach in this way.

The question that really arises in this connection is: Can we find out what the complicated computation techniques are, that are used by a particularly gifted person in this area, so that they can be made public and available to general use? If only a Margaret Singer can use techniques of that kind, then it is not a very useful kind of procedure, and I would pass on to other techniques. But perhaps with the methods she has developed and the descriptions she has put down, other people may now be able to use her methods. She has trained graduate students, and other students, I understand.

At NIMH we have other material, which has not been referred to, in relation to twins and other studies, that points to the same kind of striking and productive result. There have, of course, been many studies that were most unsatisfactory in the use of Rorschach, TAT, and so on. But I think we are getting to a place where we may be able to use these devices more effectively.

I might say to Dr. Zubin that I am not quite sure about the culture-free aspects of the responses within 1,000 milliseconds. But I do know that repeatedly in our studies we found certain differences in preparatory intervals. You used, for instance, the shift in modality. We have used different preparatory intervals, and we found the same kind of general result. We have intended to explain the results in the same way. It is a fundamental difference between schizophrenics and others. So I think with some of the leads we have, we may be beginning to get significant relationships.

If this is so, then we are getting phenomena at this very simple level which differentiate at a much more complicated level.

It is possible, of course, that we are beginning to tackle the significant problems in this area now, because we have thrown out all methods which have not turned out to be useful. But, at the same time, I would not say those devices we have thrown out

are necessarily poor. My own guess is that the operators have been poor. The devices may still have in them possibilities that we have just been too ignorant to use.

Dr. ZUBIN. I am very grateful to Dr. Shakow for pouring oil over the waters; I think he is quite right that the fault is not always in our tests, but in our testers. First, we, too, found the Rorschach worked—if you analyze it for its content. If you regard the protocols as an interview and analyze it, you get significant relationships to other ways of looking at personality. There is no question about it. It is like any other interview. But I object to looking upon it as a test because it does not qualify as such with regard to specificity of task, reliability, and validity of results. But if it is only an interview, why hide behind the ink blot or TAT card? Why not interview directly? That is the next step.

One might argue that it is advisable to have something interposed between you and the patient so as to prevent direct confrontation. But that is questionable and good interviewers do not seem to need such cushioning.

The second point I wanted to make is that the reason why I have been driven to culture-free indicators is because I have found—even in this country, aside from trying to compare with other countries—that the responses to tests were heavily laden with socioeconomic influences.

For example, when Kurt Goldstein came to this country after having demonstrated that the sorting test differentiated brain injured from normals, he began using the test in our laboratory on schizophrenics. He found that schizophrenics, too, could not adopt the categorical attitude required for sorting. But we being, we thought, a little more sophisticated, said, “Well, Professor Goldstein, how about controls?” He said, “Bring on your controls.” Who were our controls? Our graduate students! Here we had low socioeconomic level schizophrenics controlled by our graduate students. Re-

cently Vera John, working in Harlem with middle-class Negro families, found a very interesting explanation of why we may have found differences between our schizophrenics and graduate students. She found if you go into a low socioeconomic family and show a child a picture of an apple, a pear, and a banana, and ask him for one word for the three, he cannot give it. The middle class child will say, “Fruit.” Now why is it the lower socioeconomic level children cannot do it? Not because they are prone to schizophrenia! In a lower socioeconomic family, the mother says, “Johnnie, run out and get me two bananas.” Or, “Get me three apples.” But in the middle-class family, it’s, “Johnnie, go out and get me some fruit.” “What fruit shall I get, mother?” “Oh, get bananas, pears; whatever he has.” It’s the actual practice in grouping such things that makes the difference.

So we have to go over our entire repertoire of psychological tests and eliminate this differential due to socioeconomic levels; because, especially when we come to the poverty pockets, our tests are no good. They are not trustworthy elsewhere, either, because of this very fact: We have not taken into consideration important social-cultural elements in their construction.

Dr. STEIN. I would like to comment on a combination of what Drs. Zubin and Harris have been saying. Stating the results in terms of Dr. Zubin’s curve that he presented previously: If the individual has no personality and is at that point where most of the functions that exist are intellectual functions, obviously your intellectual test will do best. But if you are going to work with individuals who still have a personality, then your personality measures may be the kinds of things that provide you with data.

I would like to make a plea for the further development of conceptual frameworks and systematic variables. These are basic requirements if meaningful typologies are to be developed and if the typologies are to be useful. At the present time not all in-

investigators in the area of personality use a standardized and generally accepted nomenclature. There are a variety of terms for characterizing individuals. Some of them, although they are part of different personality theories, overlap while others do not. However, we do not yet know where the overlap does and does not exist.

A second problem in this area is that psychologists use various methods or techniques (questionnaires, projective tests, situational tests, etc.) to study the same variable. To date there has been insufficient study of

the meaning of the discrepancies which might be obtained. For example, a subject can be high on need achievement when measured by a self-description questionnaire but low on the same variable when studied with a projective test. We have yet to understand fully what such discrepancies mean for all personality variables.

Without further conceptual work and without further study of the operationalized definitions of the variables, the work on typologies will proceed in a limited manner and fall short of its mark.

Process and Reactive Schizophrenia: Some Conceptions and Issues¹

Norman Garmezy, Ph. D.²

Origins of the Process-Reactive Concept

Psychiatry's historic concern with prognosis provides the base for the evolution of the process-reactive distinction in schizophrenia. The enduring giants of the past—Kraepelin, Bleuler, Meyer, Sullivan—all grappled with the issue of prognostic efficacy and suggested factors that influence the course of the disorder; these gradually framed the dimensions that now characterize the twin concepts of process and reactive schizophrenia.

Although the separation of a process group of schizophrenics was initially made by Frank in 1932 (1), the forerunners of the distinction were provided by Karl Jaspers (2) and Eugen Bleuler (3), both of whom sought to distinguish between process and reactive psychoses.³ But the most meaningful origins of the concept must begin, as do so many aspects of psychiatric thought, with the contributions of Kraepelin.

When at the close of the last century Kraepelin first brought order to the mass of symptoms that he found in mental patients, he chose general paralysis as the disease model for creating a schema for psychiatric classification; the pattern of that model de-

creed that each mental disease be classified in terms of a specific etiology, a determinate course, and an inevitable outcome. As Zilboorg and Henry (4) point out, the distinguishing feature of the Kraepelinian system of classification (5) was a prognostic attitude in which the validity of diagnosis was determined by outcome. Thus, an appropriate outcome designated an appropriate earlier diagnosis; an unanticipated outcome reflected a diagnostic error. To this bias was added an even less viable one that continues to exert an influence on current psychiatric thought. Mental disorders were divided on the basis of exogenous or endogenous factors. Since the former presumably arose out of external events, they were potentially curable; the latter, however, stemming from inherent biological/constitutional defects, were deemed incurable. The course and the outcome of the disorder were thus predetermined and any reference to the antecedent and consequent of a morbid process was to be made in terms of the natural progression of events.

Referring to the effects of such a confluence of diagnosis and prognosis, Zilboorg and Henry note the unfortunate results that ensued:

One cannot say that because a disease ends in a certain definite way it is a certain definite disease. Kraepelin himself was apparently unaware of this singular deviation from medical principles and did not foresee that the fatalism with which it was imbued

¹ The preparation of this paper was supported by USPHS Research Grant MH-06170, a USPHS Research Career Award, MH-K6-14914, and a grant provided by the Graduate School of the University of Minnesota.

² Director, Center for Personality Research, Department of Psychology, University of Minnesota, Minneapolis, Minn.

³ The author is grateful to Dr. Robert Cancro for comments regarding the early usage of a process-reactive terminology.

weakened even further the rather unstable and never too strong rational therapeutic interest with regard to mental diseases. There is no doubt that it was not Kraepelin's intention to diminish the therapeutic efforts or to keep them only within the limits of the aging tradition of hospital management and humanitarian tolerance. But the therapeutic efforts were to become based on the complacent, expectant attitude that if the disease is a manic-depressive psychosis the patient will get well, and if it is a dementia praecox the patient will deteriorate—or, in the more turgid language of psychiatric formalism, if it is a manic-depressive psychosis "the prognosis is good for the attack," and if it is a dementia praecox the prognosis is unfavorable (4, p. 456).

Even in Kraepelin's era such a simple resolution did not coincide with factual data. In the clinics of Germany a wide range of non-deteriorated remissions in schizophrenia were noted ranging from a low of 8 percent to a high of 52 percent. Kraepelin acknowledged that a complete and lasting cure could be achieved in some instances, since his own observations and statistical studies indicated that improvements extending from months to decades were achieved in some 26 percent of his diagnosed dementia praecox cases. Here too heterogeneity was present, recovery tending to vary with diagnostic subtype; hebephrenics showed fewer favorable outcomes (some 8 percent recovering); some catatonics, on the other hand, showed demonstrably more remissions (approximately 13 percent). For Kraepelin, however, data such as these did not necessitate modification of his conceptual position. First, there was the likelihood that diagnostic errors had been committed. Second, there remained the reality that the criteria for recovery were often vague and fragile and not readily replicable from investigator to investigator. Finally, even in those instances in which improvements extended over a decade, recovery was not necessarily the most likely outcome (6).

Meyer, uncommitted to the inevitability of deterioration in schizophrenia, observed that an uncompromising theory could not be constructed on compromised data. He concluded—

that between dementia praecox and manic-depressive insanity and simple psychopathy there is an uncertain territory which refuses a categorical arrangement in the easy and simple dogmatic terms that "some disorders *must* be a deteriorative brain-disease because they early present certain signs also seen in actually accomplished deteriorations" and the claim that it would be futile to make an effort to analyze the data as a whole in terms of cause and effect (7, p. 395).

It is interesting to note that the Kraepelinian tendency to confirm diagnosis by prognosis and to revise classification, if necessary, on the basis of outcome is also held by some contemporary investigators. Langfeldt (8), for one, distinguishes between the prognostically unfavorable cases of process schizophrenia and the more prognostically favorable instances which he prefers to term "schizophreniform" psychoses. However, Langfeldt does not hold that prognosis determines diagnostic judgment but rather that the presence of certain process symptoms heightens the probability of ultimate schizophrenic deterioration. Here too the presumed relationship between process symptoms and poor prognosis proves to be something less than invariant when put to actual empirical test (9).

Still another bias provided by Kraepelin was related to his view of the personality attributes of patients during the premorbid period. Many of the traits that characterized his patients were those that we now identify with the shut-in personality—excessive shyness, withdrawn asociality, capricious demandingness. These dispositions were to be found prior to the adult life of the patient, providing a continuity with the past extending back into the patient's early years. But

again data such as these were sacrificed to theory. Kraepelin saw these behavioral patterns not as antecedents to symptom expression but as a product of the malady and its cause. What then of those persons with similar predispositions who failed to develop a schizophrenic disorder? For Kraepelin the explanation could be twofold: (1) other factors produced schizophrenic-like peculiarities, and (2) an arrested state of dementia praecox ("latent schizophrenia") could exist (6). Thus the predetermined course of a dementia praecox was reflected in the longitudinal emphasis assigned to the disorder.

By contrast, Meyer perceived the premorbid period as characteristic of "factors that are apt to shape or undo a life—specific defects or disorders of balance, with special tendencies and habitual ways of bungling and substitutions and a special make-up which is liable to breakdown in specific manners." Thus for Meyer the premorbid period was one in which complete action grew more and more disorganized by initially trivial and noninjurious subterfuges but which "in some individuals, lead further, become harmful and uncontrollable, tend to assume types of definite anomalous mechanisms, unintelligible and crazy if viewed apart, but more or less intelligible, as a string of actions substituting for, and often missing, an efficient adjustment to concrete and actual difficulties."

That Meyer identified variants within the disorder is clearly apparent from his writings:

* * * the fluctuations observable in dementia praecox are decidedly too often accounted for by renewed upsets and tangles and irritation of idiosyncracies, and that the prognosis of the ultimate tendency is remarkably often foretold, so that of the cases interpreted as actual deteriorations but few surprise us with a recovery, and those that *do* recover are as a rule specified at the outset as cases merely akin to this group worth naming by the end-stage, but

with varying amounts of balancing material (7, pp. 397–398).

It is interesting to note the consequences that follow from a Kraepelinian as opposed to a Meyerian position. The former provides its greatest contribution to a descriptive classificatory schema; the latter makes possible the introduction of a more therapeutic and preventive orientation in psychiatry. Meyer found gratifying Freud's and Jung's view that at work in the disorder was a group of complexes consisting of "insufficiently balanced experiences in various ways modified by symbolism," but he rejected an interpretation of the origins of the disorder in terms of hereditary or toxic factors.

* * * I would prefer to adhere in my attempt to define the responsible factors as far as possible in terms of prophylactic suggestiveness, in terms of untimely evocation of instincts and longings * * * and ensuing *habit-conflicts* with their effects on the balance of the person, and on the sum total of mental metabolism and actual doings and on the capacity for regulations in emergencies (7, p. 399).

Thus, despite the acknowledged difficulty of providing therapy for schizophrenic patients, Meyer's views were more positive toward amelioration:

The best procedure is to tide over the acute tangle with as much tact and ease as possible, to promote relaxation, and to relieve the situation wherever that can be done, bearing in mind the facts obtained referring to the upsetting factors, the probable complex-constellations and prevailing physical disorders. As soon as the patients feel that they meet with help instead of an argumentative and corrective attitude they can be led considerably when the time comes or where the difficulty has not led to complete blocking. Then a positive re-education in the form of habit-training and of readjustment has to set in. It is obvious that experience brings a certain divination and that individual capacity plays a decided role in the straightening out of

the difficulties, both during the tangles and in ultimately marshalling the forces to a more practical unity and level again; it is also obvious that we cannot be very optimistic in most cases, as little as when we try to win over our less unbalanced neighbors to a better mode of thought, belief, and conduct and behavior (7, p. 398).

Any brief citation of the theorists responsible for the systematization of the process-reactive concept must include an acknowledgment of the genius of Bleuler. His view that dementia praecox was not a viable disease entity but rather a group of reactions, some of which did not eventuate in deterioration, proved to be the forerunner of studies of remission and the forms of premorbid patterns, mental status factors and symptom attributes that accompanied recovery as opposed to deterioration. The influence of psychological factors was acknowledged by Bleuler who saw in secondary adaptive symptoms, such as autisms and delusions, a dependence upon psychic influences and interests. Bleuler observed that the disorder "may take a course which is both qualitatively and temporally rather irregular [with] constant advances, halts, recrudescences or remissions possible at any time"—a view that served as the pathway for subsequent studies of prognosis in schizophrenia. For if there was little variation in the course of the disorder, as Kraepelin had suggested, why study outcome at all? If, however, prognosis was irregular and not readily predictable, then premorbid life history patterns, behavioral antecedents, and symptoms could all be evaluated as predictors of the directions the disorder could take. In Meyer's words, attention would then have to be directed toward "factors which we *see at work* in the life history of cases of so-called dementia praecox."

Finally, we must turn to Harry Stack Sullivan as the exemplar of a psychiatry that emphasized factors which are to be found in the life history of schizophrenic patients.

Sullivan's contributions to the study of schizophrenia are far too numerous to elaborate, but several are particularly relevant to the study of the process-reactive concept. Far more than any other investigator, Sullivan's psychodynamic formulations and his great interest in the psychological antecedents and correlates of a schizophrenic disorder reflected the growing preoccupation in America with a more developmental and psychologically oriented view of the disease process. In brief, this viewpoint had several foci: (1) an increased emphasis upon life history antecedents and the role of precipitating stresses in actuating the disorder; (2) a view of the schizophrenic process itself as a reintegrative effort at adaptation; (3) a rejection of the prevailing dogma that recovery, by definition, implied a misdiagnosis of schizophrenia; (4) emphasis on the role of the hospital and its milieu in blocking or heightening chronicity; (5) the role of sexual and social adaptation in the disorder; (6) a stress on interpersonal factors and cultural distortions provided by the home with particular emphasis on the "subject's life experience as experienced—i.e., what he actually lived or underwent, not what he is alleged to have experienced"; (7) the perception of stress as arising from societal relations, rather than from impersonal physical factors; (8) the significance placed upon a "really satisfying adjustment to a sex object" as a markedly favorable prognostic sign; (9) rejection of a simple dichotomous organic-psychogenic etiology; (10) the differentiation of social (non-symptomatic) versus personal recovery (i.e., reorganization of the disordered personality); and finally, (11) the importance of psychological intervention as a requisite for a favorable outcome in some cases.

The flavor of Sullivan's views can be captured by many quotations from his writings. The following illustrate his biases with regard to the issues of etiology, course, and outcome in schizophrenia:

Antecedents

Study of onset of disorder in male patients * * * seems to establish two factors preliminary to schizophrenic psychoses. Firstly, the appearance of the disorder is late in a long series of subjectively difficult adjustive efforts. Secondly, it seems never to occur in those who have achieved if only for a short time a definitely satisfying adjustment to a sex object (p. 104).

One finds that the individual who has had a schizophrenic illness has not, in the first place, developed the abrupt manifestations of hereditarily-determined deterioration in the life processes. Instead, he has stood in a significantly and distinctively difficult position in the social situation in which he has lived; he has developed a striking, more or less specifically distinct technique in dealing with people with whom he has lived; in the course of this peculiarly distinguished life he has come upon certain situations which were most serious in their negative effect upon his self-esteem; and after encountering these situations (which include as significant factors only other people) after, perhaps, a rebuff to his self-assertion, he has shown a significant and characterizable failure to react by any of the methods of reacting to rebuff which are more or less well known to all of us from our personal experience * * * (10, p. 221).

Etiology

It is easy to divide the material under consideration into "praecox" illnesses based on organic pathology and "schizophrenic" illnesses based on functional pathology. The division, however, is irrational and unprofitable, for some of the former cases recorded a good measure of mental health just as did most of the latter. In other words, in frankly defective patients, undergoing severe and relatively typical schizophrenic processes, nothing fully distinctive from extraordinarily talented individuals suffering schizophrenia has appeared in this investigation (p. 240).

Historically, psychiatry has been a field peculiarly afflicted by bad thinking and premature hypothetic formulation. Of many reasons for this, I will touch

upon two. The facts of mental illness have often been seen through the aberrating medium of patho-physiological preconceptions—brain pathology and the like, endocrine disorders being perhaps the most recent. And the physician has been carried into the recondite field of mind without training in the technologies suiting him to his enterprise, and with training tending definitely to disqualify him for perceiving the data on which he should base his conclusions. Especially in the field of the schizophrenic phenomena, there has been a pandemic of formulating on limited if not actually irrelevant basic data, with singularly bad hypotheses, some of the most vicious features of which have been incorporated into persisting psychiatric astigmatisms (10, p. 321).

Outcome

* * * it is the life situation of the patient that determines the prognosis. What he has derived from his forebears, his life experience, and that which befalls him during his illness—these, correlated with the situation which confronts or seems to confront him in the event of his recovery—these, and these only, are the determining factors which make in their biological summation, for benignity or malignancy of the situation (p. 27).

Psychiatric prognosis may best be considered as a specialized technique in social psychology. Its problem is the prediction of the future adaptability of an individual within some more or less clearly envisaged milieu composed principally of people. To reach a judgment of prognosis, facts are accumulated in regard of (a) the personality of the patient, (b) the morbid process which he is suffering, and its effects on his personality, and (c) the significant factors in that cultural environment to which the patient may presently be returned. In no case an easy achievement, prognosis of those individuals who suffer schizophrenic disorders is peculiarly difficult for several reasons. The personality of the patient is hard to appraise: his cooperation in such procedures as free association and dream study are difficult to secure; facts contributed by previous associates are

meagre, owing to his shut-in responses to pre-psychotic social situations; and the parents are almost always psychopathic in their reaction to investigation—perceiving more or less dimly their part in the illness of their progeny and smoothing over and falsifying liberally. Again, the fundamental characteristics of the schizophrenic process itself are still in doubt. We are but beginning to free ourselves of many misconceptions concerning “dementia praecox,” and the majority still revise diagnosis when the patient recovers. The Kraepelinian diagnosis by outcome has been a great handicap, leading to much retrospective distortion of data, instead of too careful observation and induction. Finally, it is no small task to distinguish in the environment to which the patient may expect to return, those cultural factors which are potentially effective in his adjustment. Without accurate data as to his own valuations, we are prone to serious errors in deciding which factors are the significant ones and our efforts to weigh their influence are relatively futile (10, pp. 158–159).

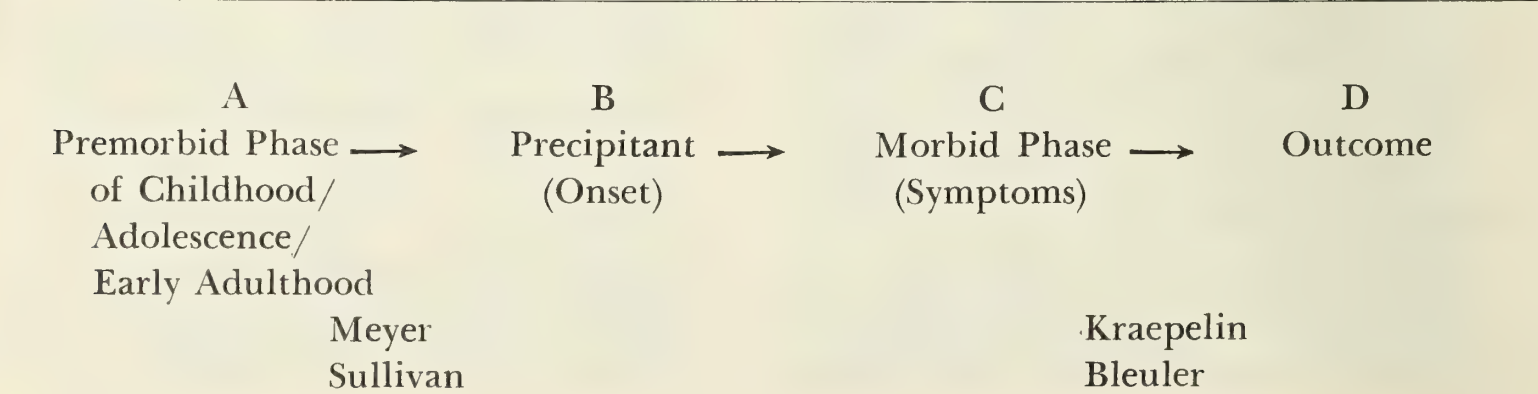
“group of schizophrenias” in which the likelihood of remission and recovery were probabilistic events opened the path to a review of the role of psychic influences in the disorder. Meyer and Sullivan brought these influences further into psychiatry, emphasizing a developmental tradition that was characteristically prevalent in American psychology—one in which life history antecedents and prepotent habit patterns were viewed as precursors to the disorder and markedly relevant to outcome.

If one were to mark a developmental trail from the premorbid period to the ultimate outcome in schizophrenia, the significant prognostic points in the sequence would be those charted (A-B-C-D) below.

To summarize their prognostic contributions, Kraepelin and Bleuler, wedded to biological causation, accentuated prognostic criteria derived at point C in the sequence, i.e., symptoms as predictors of outcome. The more psychologically-minded viewpoints of Meyer and Sullivan were reflected in the emphasis given to prognostic contributions at points A and B in the sequence.

One can perceive elements of this lineage in an experimentally-minded present. The parallel to C → D type studies is revealed in factor-analytic investigations of symptom patterns and outcome; patterns of physiological responsivity in schizophrenia; psychometric test studies, etc. By contrast, the A/B → D relationship often takes the form of the manipulation of cue relevant variables (e.g., stress,

In summary, the concepts of process and reactive schizophrenia arose from a fundamental concern with prognosis. This interest in outcome, however, stemmed initially not from a therapeutic-rehabilitative orientation but rather from a theoretical-diagnostic base in which diagnosis and prognosis were invariant. Exceptions to (presumed) diagnosis-prognosis lawfulness were treated as the consequences of diagnostic error and faulty criteria for recovery rather than instances that warranted conceptual review. Bleuler’s conception of a



censure, symbolic representations of familial interactions, portrayals of interpersonal and sexual relationships, etc.) in conjunction with the study of schizophrenic adaptation or disorganization.

Attributes of the Process-Reactive Dimension

The process-reactive concept incorporates all points in the developmental sequence of the schizophrenic disorder since wide variations can be observed among patients at each point. Specifically, variation in the premorbid personality patterns of patients, the time interval preceding the onset of the disorder, the type of precipitant that may potentiate the disease process, and the symptom picture during the course of the disorder all have been related to prognosis in schizophrenia.

Although the concept of a process as opposed to a reactive schizophrenia has the widest acceptance, other, but comparable, dichotomous terms continue to be used—*dementia praecox/schizophrenia*, *chronic/episodic*, *typical/atypical*, *true/schizophreniform*, *evolutionary/reactive*—to reflect the varying patterns present in the schizophrenic disorder.

In general, the picture presented by a so-called process case may be described as follows: The patient's prepsychotic personality is a poorly integrated one revealing markedly inadequate behavior in the sexual, social, and occupational areas; trends to social isolation and a lack of emotional responsivity to others are clearly evident. There is usually no acute precipitant to characterize the turn toward psychosis; rather, the onset (usually in late adolescence) is an insidious one without a recognizable and consensually validated stressor evident. Symptomatically, there is a gradual onset of emotional blunting, a withdrawal from life's daily activity; apathy and indifference hold sway and somatic delusions and marked disturbances in thinking may characteristically be pres-

ent and maintained for long periods of time.

By contrast, the reactive patient has been described by Wiener as follows:

From birth to the fifth year, the maturational and developmental history showed no defects, physical health was good. Generally school and home adjustment was good. Parents were accepting. Heterosexual relationships were established. The patient had friends, and domestic troubles did not disrupt his behavior.

The onset of the illness was often sudden with a clear-cut, understandable precipitating event. Aggression was expressed verbally. Decency was retained. The course was fulminating, with massive hallucinatory experiences, ideas of reference, and mild paranoid trends, as well as sensorial impairment. A thought disorder was present according to some authors, but not others. Response to treatment was good (11, pp. 156-157).

Kantor, Wallner and Winder have provided an excellent summary of the criteria that best differentiate the process case from the reactive one. (12) As indicated in table 1, these criteria embrace all points in the developmental sequence of the disorder.

The link between the case history criteria in process and reactive schizophrenia provided by Kantor et al. and favorable and unfavorable prognostic indices as formulated by Huston and Pepernik (13) can be seen by referring to table 2. There is a scattered amount of surplus content in both listings. For example, body-type is a rather out-moded variable that does not claim the attention of contemporary investigators. Similarly, the severity of physical illness in the earliest years lacks strong empirical validation. Other variables such as the form of manifest aggression, IQ level, early psychological trauma versus the presence of a good psychological history and the capacity for alcohol in adulthood have not been sufficiently researched to command attention. However, other variables have

TABLE 1.—Case History Criteria for Differentiating Process and Reactive Schizophrenia
[From Kantor, Wallner, and Winder]

Process schizophrenia	Reactive schizophrenia
BIRTH TO THE FIFTH YEAR	
a. Early psychological trauma.	a. Good psychological history.
b. Physical illness, severe or long.	b. Good physical health.
c. Odd member of family.	c. Normal member of family.
FIFTH YEAR TO ADOLESCENCE	
a. Difficulties at school.	a. Well adjusted at school.
b. Family troubles paralleled with sudden changes in patient's behavior.	b. Domestic troubles unaccompanied by behavior disruptions. Patient had "what it took."
c. Introverted behavior trends and interests.	c. Extroverted behavior trends and interests.
d. History of breakdown of social, physical, mental functioning.	d. History of adequate social, physical, mental functioning.
e. Pathological siblings.	e. Normal siblings.
f. Overprotective or rejecting mother. "Momism."	f. Normally protective, accepting mother.
g. Rejecting father.	g. Accepting father.
ADOLESCENCE TO ADULTHOOD	
a. Lack of heterosexuality.	a. Heterosexual behavior.
b. Insidious, gradual onset of psychosis without pertinent stress.	b. Sudden onset of psychosis; stress present and pertinent; later onset.
c. Physical aggression.	c. Verbal aggression.
d. Poor response to treatment.	d. Good response to treatment.
e. Lengthy stay in hospital.	e. Short course in hospital.
ADULTHOOD	
a. Massive paranoia.	a. Minor paranoid trends.
b. Little capacity for alcohol.	b. Much capacity for alcohol.
c. No manic-depressive component.	c. Presence of manic-depressive component.
d. Failure under adversity.	d. Success despite adversity.
e. Discrepancy between ability and achievement.	e. Harmony between ability and achievement.
f. Awareness of change in self.	f. No sensation of change.
g. Somatic delusions.	g. Absence of somatic delusions.
h. Clash between culture and environment.	h. Harmony between culture and environment.
i. Loss of decency (nudity, public masturbation, etc.).	i. Retention of decency.

sturdy empirical supports and these factors, in the form of dimensionalized attributes, are included in the two scales that are most frequently used as indices for process and reactive status: the Elgin Prognostic Scale, (14, 15) and the Phillips Scale of Premorbid Adjustment in Schizophrenia (16). Both scales have the marked virtue of supplanting vague and subjective descriptions of unscaled dimensions with quantitatively ordered criteria that bear a demonstrably reliable relationship to prognosis in schizophrenia.

Rating Scales for Process-Reactive Schizophrenia

Elgin Prognostic Scale

The original Elgin Prognostic Scale was devised by Wittman on the basis of a review of some 50 studies of prognosis in

schizophrenia. Originally, it was comprised of 30 subscales (25 measured premorbid adjustment and five the presenting symptoms) that were subsequently trimmed to 20; each scale carried "armchair" weights that reflected the prognostic significance of the items based upon clinical judgment. Factors at the favorable end of the scale were arbitrarily assigned negative values, those at the unfavorable pole positive values. An algebraic sum of these various weighted measures constituted the overall prognostic index. Case history data was the preferred locus for the ratings. The various scale items included those factors related to prognosis to which reference has been made previously: childhood personality patterns, rate of onset, precipitating events, presenting symptoms, duration of

TABLE 2.—Favorable and Unfavorable Prognostic Indices in Schizophrenia
[From Huston and Pepernik]

Favorable	Unfavorable
CLINICAL FACTORS	
Acute onset*	Gradual onset.*
External precipitant*	
Short duration of illness prior to hospitalization.*	Long duration of illness prior to hospitalization.*
Preservation of affect.*	Inappropriate affect.*
	Flat affect.
Tension and anxiety.	
Depression.	
Manifest moderate hostility.	Little overt hostility.
Self-reproaching delusions.	
Compliantly paranoid.	Defiantly paranoid.
Ability to rationalize reality lapses.	
SOCIAL AND PERSONAL HISTORY FACTORS	
Upper socio-economic group.	Lowest socio-economic group.
Good educational history.	
Good occupational history.	
Steady church attendance.	
Married.	Single, divorced or separated.
Good marital adjustment.	
Good recent sexual adjustment.	
Relatively stable prior to onset.	
Under 30.	Over 35.
PSYCHOLOGICAL TEST FACTORS	
High IQ.*	Low IQ (under 90).*
Functioning at lower level than potential.	
Little impairment of abstract thinking.	Marked impairment of abstract thinking.
Favorable Rorschach indices (M, C, total R, F).	
CONSTITUTIONAL FACTORS	
Ecto-mesomorphic.	Endomorphic.
Pyknic.	
PSYSIOLOGICAL FACTORS	
Chills after mecholyl.	
Anxiety precipitated only by mecholyl.	
Small blood pressure response (or decrease) to mecholyl.	Increased blood pressure response to mecholyl.
Moderate blood pressure response to epinephrine.	

NOTE: asterisked items reflect more reliable prognostic indices

psychosis, and body-build. Although each scale had several rating points, only the polarities bore descriptive statements. Reliability reports on the scale are surprisingly few. McDonough (17) cites an interrater reliability of +0.89 obtained from independent readings of case history data. This is consistent with Wittman's two-rater reliability of +0.87. Chapman, Day, and Burstein (18) report agreement of 0.97 between two raters after extensive resolution of differences in interpretation of items and ratings accorded these items by the clinical judges. (Other

studies have used the criteria set forth by Kantor, Wallner, and Winder with reasonably high degrees of reliability but these studies are not relevant to the reliability of ratings provided by the EPS (19, 20, 21). Regarding the components of the scale, these are best revealed in a factor-analytic study conducted by Lorr, Wittman, and Schanberger (22). Using Elgin scale ratings provided by two clinical judges on 200 successive admissions of hospitalized psychotics, the factor analysis revealed three factors: schizoid withdrawal; schizophrenic reality distortion; and a somewhat poorly

defined factor of personality rigidity and maladaptation.

On the first factor, 13 of the 20 scales have factor weights of 0.35 or greater. Typical of such high loadings are scales tapping defects of interest, insidious onset, shut-in personality, schizothymic personality, limited range of interests, low energy tone, lack of heterosexual contacts, etc. As Becker (23) has pointed out, the first factor loads so heavily on relevant dimensions that it alone probably reflects the elements that constitute the process-reactive dimension. The overweighting of this first withdrawal factor is seen in the subscales that merely reflect minor variants of the same behavior. The second factor, reality distortion, has also been confirmed by Becker and this together with the emotional rigidity reflected in Elgin ratings of traits characterized as careless indifference and exclusiveness-stubbornness enables a judgment of severity of pathology to be made. Poor prognosis and greater severity of disorder are reflected in high scores on items comprising the factors; a more favorable prognosis typical of the reactive schizophrenic would be characterized by low scores on the subscales.

Since the appearance of the Elgin scale, Becker (15, 24) has created a revision that provides for more precisely described intermediate points within each subscale thus strengthening the likelihood of more reliable ratings by clinical judges. This revised scale is presented in appendix A. Subsequently, Steffy and Becker (25, 26) created an abbreviated version of the Elgin Scale on the basis of the factor analytic research that has been cited. Those items that constitute the revised form are indicated by an asterisk.⁴

The validity of the original scale was successfully determined initially on the responses of 343 diagnosed schizophrenics to various types of shock treatment (27). The scale was far more effective in predicting outcome results than were the prognostic estimates made by staff members. Subsequently, Wittman reported two additional studies (28), one conducted with 825 schizophrenic patients and another with 804 schizophrenic and 156 manic-depressive patients again comparing the predictive efficacy of scale versus staff on clinical status of the patients following various forms of treatment. In these studies also the EPS proved to be superior to the clinical judgments of staff members.

Attesting to the validity of the Becker revision of the Elgin Scale is the more recent study of prognosis by Chapman et al. (18). These investigators used a mental hygiene clinic as the base for their study, since it provided for a brief psychiatric stay before patients were either discharged or referred to one of seven State hospitals. Subjects in the study included an interracial group of 106 diagnosed male schizophrenics. Of this group approximately two-thirds were first admissions, the remainder having had one or two prior admissions totaling, on the average, 5 months of hospitalization. EPS scores were obtained on the basis of social history data and supplemented by interviews in order to rate those subscales related to presenting symptoms.

Subsequent follow-up permitted biserial correlations to be computed for EPS scores and patients' discharge status for 1- to 9-month periods following admission. It was not until the sixth month after admission that the scale began to predict discharge reliably, but even at 9 months the correlation ($r_{bis} = +0.38$) predicted less than 16 percent of the variance. Furthermore only 11 of 20 items of the scale were significant predictors of discharge status with no scale exceeding an r_{bis} of $+0.37$. Thus, the EPS predicted discharge only slightly

⁴ This abbreviated version includes the following subscales: I(A) Defects of Interest versus Definite Display of Interest; II(B) Insidious versus Acute Onset; III(C) Shut-In Personality; IV(I) Heterosexual Contact; V(K) Careless Indifference versus Worrying, Self-Conscious Type; VI(L) Exclusive Stubborn Traits versus Insecurity and Inferiority Feelings; VII(N) Precipitating Conditions (Situational Reaction); VIII(D) Duration of Psychosis; IX(Q) Hebephrenic Symptoms; X(S) Physical Interpretation of Delusions.

better than chance and no better than knowledge of the patient's marital status alone. However, appropriate consideration must be given to the many factors that can lead to a summary discharge of patients from a mental hospital. It is quite likely that either a more adequate criterion of improvement or a more stringent criterion for recovery (e.g., a ratio of time out to time in the hospital over a prolonged period of time) would demonstrate more reliably the prognostic utility of the EPS. Data to be cited shortly regarding the prognostic utility of the Phillips Scale attest to the need for a more rigorous criterion of recovery than the one employed by Chapman et al.

Of considerable interest is the observation by these investigators of the difficulties encountered in using the EPS. I take the liberty of quoting extensively from their appraisal since these authors have provided a rather definitive critique of the instrument:

The wording of the descriptive labels which Wittman offered for items, although clinically rich and descriptive, is exceedingly imprecise. It is not entirely clear what is meant by "shut-in" personality or "poor bite on life." Even phrases like "withdrawal and disinterest in social surroundings" or "dates frequently" were found to be decidedly ambiguous in the absence of specification of the degree and variety of withdrawal, or how many dates. Ambiguity in some items results from a lack of specification of whether the material is to be rated at a behavioral or an interpretive level—e.g., "insecurity and inferiority feelings" (item L). Also, some of the items call for a judgement which is very difficult to rate objectively, even with optimal information—e.g., "low-energy tone" (item G).

Furthermore, in at least three items, two characteristics are stated as opposites which are not necessarily opposite. For example, in item J the rater must score "marked academic interests versus active interests in sports"; in item K he must score "careless indifference versus worrying, self-conscious type"; in item L he must score

"exclusive stubborn traits versus insecurity and inferiority feelings." In the present sample several patients were both studious and athletic, or both self-conscious and carelessly indifferent, or stubborn with feelings of inferiority. Also, some patients showed neither of the two characteristics—e.g., they were neither studious nor athletic.

Many single items contain a variety of characteristics not necessarily related, with the requirement that they all be rated on a single continuum. For example, we are asked to rate in item O:

"Shut-in personality: General.—The psychotic condition is simply an exaggeration of the peculiar type of personality shown all through childhood. Stormy childhood often with overprotection and anxiety, a difficult adolescence characterized by inability to get along with and mix with other children. Constitutional apparently rather than product of specific environment."

In addition many of the items appear to refer, at least in part, to the same variables. The descriptive paragraphs of nine of the 20 items mention, among other things, something about inadequate social interests * * *

There are also items on the EPS which might be expected to have differing implications for level of adjustment, depending upon the subculture; e.g., item E, "range of interests"; item J, "marked academic interests versus active interests in sports"; item K, "careless indifference versus worrying, self-conscious type." A well-adjusted young man of upper middle-class background would more likely have wide interests (a reactive sign) and academic interests (a process sign) than would a well-adjusted member of a lower class group. Also, "careless indifference" is probably more socially acceptable in some lower class groups than in the middle class (18, pp. 387–389).

To these disadvantages must be added others. The EPS is an unwieldy instrument, far too elaborate, too multivariate, too inclusive of scales that fail to correlate with prognostic criteria, too biased in favor of an old-fashioned constitutional and somatotypic base, and, above all, dependent upon life history and mental status data far

more extensive than those usually available in the typical case history file. Becker's abbreviated version of the scale offers many improvements but some of the shortcomings remain.

The Phillips Scale of Premorbid Adjustment in Schizophrenia

In 1953 Phillips reported the results of an outcome study using a newly devised ordinal prognostic rating scale (16). The scale was originally developed (and subsequently cross-validated) using case histories of schizophrenic patients who had improved or failed to improve following some form of shock treatment. Initially the scale consisted of three subsections: (a) Premorbid History; (b) Possible Precipitating Factors; and (c) Signs of the Disorder. With each subscale there were further subdivisions containing brief descriptive statements with empirically (and, in part, intuitively) determined weights. Each item of the subscales ranged either from 0 or 1 to 5 or 6.

In creating the original subscales, a score of 3 was assigned to divide the improved from the unimproved groups of patients. The smaller the weighted score, the more significant was the item as an index of improvement; the greater the weighted score, the more definitive the index of non-improvement.

The final order of weights for the items of each subscale was determined by those case history findings that characterized the two preliminary groups of patients Phillips had used in his initial study. In some instances it was possible to create items with seven steps representing increasing adequacy of adjustment. In other instances, only four such points could be derived from the data. Two other groups of patients were then rated on the scale to check its validity as a prognostic index. The final validating group was comprised of 31 patients, 29 of whom had been diagnosed schizophrenic, with the remaining two as paranoid conditions. Five clinical judges rated these pa-

tients 6 to 12 months after some form of shock treatment (electroshock, insulin coma, metrazol, or a combined treatment) as: "greatly improved," "moderately improved," and "least improved."

Phillips found that case data related to the recent premorbid period bore the most unequivocal relation to outcome, particularly that subscale which dealt with the social aspects of the patient's recent sexual life. Recent sexual adjustment and the past and present adjustment in interpersonal relations also proved to be particularly effective. By contrast there existed only a suggestive relationship between the precipitating event and outcome (a finding supported by Chapman, Day, and Burstein's analysis of the comparable EPS items). Data derived from signs of the disorder suggested that affective reactivity and the degree of mood and thought disorder also bore a relationship to therapeutic potential. This last group of items, however, is highly dependent upon the presence of excellent mental status accounts in the case folder. Since Phillips had found a marked tendency for premorbid history scores to correlate highly with signs of the disorder⁵ ($r=+0.91$) and to a lesser extent with possible precipitating factors ($r=+0.70$), other investigators turned to the premorbid history subscale as the sole criterion for assignment of schizophrenic patients to disparate prognostic groups (30). This scale (now termed the "Phillips Scale of Premorbid Adjustment in Schizophrenia" or, more simply, "the Phillips Scale") has a number of advantages over the Elgin Scale. First, it avoids such elusive concepts as constitutional bias, low energy tone, asthenic build, toxicity of exhaustion. Second, it demands only minimal case history data to secure reliable patient ratings. Third, the reliability of the scale has been vigorously substantiated. Fourth, its construct validity has been elaborated

⁵ Seidel (29) reports a correlation between the premorbid history scale and total prognostic scale of $+0.96$.

through a series of interdependent and independent studies. On the other hand, the scale has a number of short-comings as a psychometric instrument, including inadequately defined variables, the assignment of somewhat arbitrary weights, etc.

The Premorbid Scale consists of five subscales: Recent Sexual Adjustment; Social Aspects of Sexual Life during Adolescence and Immediately Beyond; Social Aspects of Recent Sexual Life (with separate criteria for those above and below 30 years of age); History of Personal Relations; and Recent Adjustment in Personal Relations.⁶ Since each subscale has a range of weighted values from 0–6, the possible total range of scores extends from 0–30. Recalling Phillips' intention to establish a subscale value of 3 as the fulcrum point between improved and unimproved cases, scores from 0–15 have been used to denote so-called good premorbid schizophrenic cases, with 16–30 representing the range for assignment as a poor premorbid case. More recently, there has been a move on the part of investigators to remove the ambiguity posed by mid-scores and to set the scale criteria for designating good and poor premorbid cases as 12 and below and 17 or 18 and above, respectively.

Reliability Data

A variety of reliability studies have been conducted with the Phillips Scale (31). These studies have included variations both in the source of the rated data (patient interview, parent interview, and case history files) and the degree of clinical sophistication of the raters. Typical of the findings are those presented in table 3.

The data of table 3 clearly support the contention that the Phillips Scale is a highly reliable instrument when the total score is used for assignment to a poor or good premorbid category. Individual sub-

scales vary in reliability, with the two related to the past history and present pattern of personal relations being most unreliable. Encouraging is the fact that when the patient is used as an informant, in only five of 46 cases is a category reversal obtained.⁷ Furthermore, patients tend to report their premorbid histories as similarly as do other family members. It is possible to interview cooperative patients and, on the basis of the information supplied by them in a structured interview situation, to secure a reliable assignment to either a good or poor category. An additional bit of data obtained by the Duke investigators is that marital status accounts for most of the variance in the Phillips scale, the point biserial correlation of marriage and total score being $+0.78$. Furthermore, in a study of 65 male schizophrenics, 50, or 77 percent, could be correctly classified as good or poor by a knowledge of marital status alone. Although the designation of marriage invariably provides a rating of good premorbidly, single patients (particularly youthful ones) may rate either in a good or poor premorbid category. These data on the significance of marital status as a prognostic index are substantiated by the observation by Chapman et al. that EPS scores bear a strong relationship to marital status, the married patients being classified at the reactive end of the scale and the unmarried at the process end. Of equivalent importance was the finding that marital status alone could predict discharge from the State hospital as well as did the EPS. These results coincide with those obtained in the VA psychiatric evaluation project (33).

Validity Data

The prognostic validity of the Phillips Premorbid Scale is attested to by several studies. Farina and Webb (34) employed the Premorbid Adjustment Scale to test its

⁶ A copy of the premorbid scale with guidelines for scorers as modified by A. Farina and N. Garmezy appears in appendix B.

⁷ These findings suggest that a 24-item, self-report measure of the process-reactive continuum that has been developed recently by Ullmann and Giovannoni may prove to have considerable utility in future research (32).

TABLE 3.—Abbreviated Summary of Several Reliability Studies of the Phillips Scale

[Data from Garmezy, Farina, and Rodnick (31)]

Study	Source of data and informants	Raters	Correlation	Number of category assignment errors
I.....	Interview: 17 male schizophrenic patients versus parents. Case history versus patient interview. Case history versus parent interview.	Two senior clinicians—one for patients, another for parents. Senior clinician. Senior clinician.	+0.59 +0.62 +0.80	2/17
II.....	Case history data: 13 male schizophrenic patients.	Senior clinician versus graduate student trainee.	+0.84	0/13
III.....	Case history data: 10 male schizophrenic patients.	Senior clinician versus undergraduate research assistant.	+0.90	1/10
IV.....	Case history data versus patient interview: 14 male schizophrenic patients.	Senior clinician I versus senior clinician II.	+0.92	0/14
V.....	Case history versus patient interview: 15 male schizophrenic patients.	Senior clinician I versus senior clinician II.	+0.92	
VI.....	Case history: 15 male schizophrenic patients.	2 second year clinical psychology graduate students.	+0.90	3/15
VII.....	Case history data: 80 female schizophrenic patients.	2 senior clinical psychologists.	+0.86	
VIII.*	Interview: 29 male schizophrenic patients Interview related to patient's life history using non-psychotic brothers as informants (N=26). Interview: patient and brother (N=20). Interview: patient and brother (N=18).	Senior graduate clinician and undergraduate research assistant. As above. Senior graduate clinician. Undergraduate research assistant	+0.96 +0.95 +0.98 +0.94	

* Data from W. Bradford.

efficacy with two groups of male schizophrenic patients: one group had successfully remained out of the hospital for at least 18 months (success group); another group was unable to remain out of the hospital on trial visit or discharge status for longer than a 3-month period (failure group). Success or failure of short-term trial visits bore only a slight relationship to premorbid scale scores. However, when the criterion of later hospital status was used (e.g., whether patient was in or out of the hospital 4 to 10 years later) the scale separations revealed a high level of statistical reliability. Seidel too has demonstrated a significant positive correlation ($r_{bis} + 0.46$) between premorbid scale scores for 63 male schizophrenics, 31 of whom had been hos-

pitalized continuously for at least 3 years and 32 who had been discharged in less than 3 years with a psychiatric evaluation of either "social recovery" or "recovery" (29). In another study utilizing female schizophrenic patients, Farina, Garmezy, Zalusky, and Becker (35) employed a stringent criterion of recovery. All female first admission patients to a State hospital over a 3½-year span were followed for a period of 5½ years and the time spent in and out of the hospital for each patient during that period was cumulated. Two extreme groups were then selected—a recovered group (N=50) consisting of all patients who had remained in the hospital no more than 6 months and who had been out of the hospital for at least 5 years, and a nonrecovered

group that had been hospitalized for 5 years or more and had remained out of the hospital for a total period of less than 6 months. All patients were given Phillips ratings by a clinical judge who had no knowledge of the patients' hospital histories. Both marital status and premorbid scale ratings were found to differentiate significantly between the two groups of patients. Since married patients were found in both groups, the authors proceeded to rate this select subsample in terms of the scores obtained on scales D and E (History of Past and Present Personal Relations) as a measure of social functioning. Scores on these scales of patterns of personal relationships were found to differentiate adequately between seemingly homogeneous groups of married schizophrenic women.

A subsequent and comparable study of male patients yielded similar prognostic results. (36).

Relationship Between Phillips Premorbid Scale and Elgin Prognostic Scale

Both the Phillips and the Elgin scales have had their proponents and enthusiasts. But the question as to whether or not one can incorporate data obtained from one instrument with data obtained from the other remained unanswered until the publication of a recent study by Solomon and Zlotowki (37). These investigators rated 46

case histories using the Elgin Prognostic Scale (EPS) and the Phillips Prognostic Rating Scale (PRS). Their data indicate inter-rater reliability coefficients of +0.90 and +0.92, respectively. Their most important findings were these:

a. The presenting clinical picture does not correlate markedly with the premorbid adjustment data of either the EPS (items A-N) or the PRS. These correlations are +0.35 and +0.31, respectively;

b. A high degree of relationship exists between the two scales; the overall correlation based upon total scores is +0.87; for the Premorbid Scale alone the correlations with total EPS is +0.78. These and related reliability coefficients are presented in table 4 below. These data clearly support generalizing from data obtained with either scale.

In the remainder of this paper I will concentrate on efforts to review briefly the construct validity of the instruments that extend beyond mere prognosis. No other methods for defining the process-reactive continuum will be cited although other procedures have been employed and have been well reviewed in a substantial paper by Higgins (38).

Some Issues Involving the Process-Reactive Concept

In this section I will not attempt to deal with a review of the experimental studies

TABLE 4.—Relationship Between EPS and PRS Subscales
[From Solomon and Zlotowski]

Becker (rev.) subscales						
Number items	Phillips subscales	(14) Number items premorbid adjustments (A-N)	(1) Precipitants (N)	(5) Clinical picture (P-T)	(19) Total Elgin	(13) Total Phillips
5	Premorbid adjustment (1a-e)	0.80	0.37	0.25	0.78	0.87
2	Precipitants (2a-b)	0.63	0.44	0.25	0.64	0.72
6	Clinical picture (3-4)	0.45	0.18	0.67	0.59	0.71
13	Total Phillips	0.81	0.40	0.52	0.87
19	Total Elgin	0.96	0.48	0.50

that have been performed on subjects categorized as either process-reactive or good-premorbid/poor-premorbid schizophrenics. The interested reader is referred to articles by Higgins (38), Herron (39), Rodnick and Garmezy (30), and Garmezy and Rodnick (40) for a review of such research. Instead, related research will be reviewed within the context of specific issues posed by the process-reactive (or poor premorbid-good premorbid) concept.

Biases Inherent in the Process-Reactive Concept

The history of the process-reactive concept reinforces the view of schizophrenia as a dichotomous typology influenced *either* by a presumed somatic etiology *or* a psychogenic base.

Commenting on the reasons for such a persistently-held viewpoint, Garmezy and Rodnick wrote as follows:

The fact that some types of schizophrenia seem to have a readily identifiable psychogenic precipitant whereas no obvious stressors are seen in other cases all too frequently has led to the conclusion that somatic considerations alone influence the latter group.

Discomfort with the current state of psychological research—its complexity, diffuseness, imprecision and (at times) superficiality—has led some investigators to look longingly toward the greater exactitude and lesser complexity promised by the biological sciences. Recent advances in studying the biology of schizophrenia undoubtedly have reinforced the belief that the complex problems of the genesis of this disease are more researchable in fields such as biochemistry and physiology. Coupled with these advances has been an awareness by investigators of the great methodological difficulties inherent in any effort to reconstruct retrospectively, through the verbal reports of patients or parents, the early family milieu of an individual.

The biases of individual investigators have also played a role in accentuating a simplified dichotomous conception of the disorder. Benjamin has pointed

out how the reductionistic biases of some investigators deny the possibility of attributing causality to psychological variables. Graduate training in all disciplines all too frequently produces the type of tunnel vision in research to which Benjamin has assigned the very appropriate terms of “biophobia” or “psychophobia.”

The comfort of the “either-or” solution creates ready adherents to a mind-body dualism. The lack of an adequate theoretical conceptualization of schizophrenia which can effectively incorporate both psychological and biological variables tends to foster such simple choices (40, p. 451).

There are really three issues contained in this passage: (1) Are process and reactive schizophrenia two separate and distinct disorders or do both stem from a common disease entity? (2) Is the process-reactive distinction best viewed as a dichotomous typology or as a continuum of which the process-reactive designates are merely the polarized extremes of the dimension (3)? To what extent is it valid to equate process schizophrenia with an organic etiology and reactive schizophrenia with psychological antecedents? Although determinate answers to these questions are not available, a variety of research studies indicate the more likely form such answers will take.

1. *Process-Reactive: A single or multiple disorder?*

The clue to this issue can be found in data related to symptom expression during the disorder. We have shown how the clinical symptom subscales of the EPS and PRS lag in prognostic significance (16, 18). In Becker's (24) abbreviated version of the EPS, three of the five clinical scales—“inadequate affect versus emotional instability or appropriate affect,” “ideas of influence,” and “atypical symptoms”—have been eliminated because of their prognostic limitations; only “hebephrenic symptoms” and “physical interpretation delusions” are retained. The data provided by Chapman et al. (18) suggest that of these two only “hebephrenic

symptoms" relate to discharge status nine months after admission (although affect and atypical symptoms, discarded by Becker, reveal prognostic significance). The research of Solomon and Zlotowski (37) indicates that the symptom items of the Phillips and Elgin scales have "substantially lower interrater reliability as well as item-to-total reliability when compared with the premorbid adjustment set." These latter findings may merely attest to the inherent unreliability of scales that are based on fragile mental status data rather than on any inherent similarity between the process and reactive groups during the morbid phase of a schizophrenic disorder.

Two recent unpublished studies of presenting symptoms at the time of admission to a psychiatric hospital in which the Minnesota Multiphasic Personality Inventory was used to compare the overt symptom patterns of good and poor premorbid patients are of interest.

Schaefer (41) sought to examine the question of whether schizophrenia was a valid appellation to assign to newly admitted good and poor premorbid patients. Her method for answering the question was to examine the symptom picture presented by acutely ill patients in each group at the time of their admission to the hospital, using the patients' self-reports on the MMPI to evaluate and compare symptomatology. Although the study could not possibly answer the ubiquitous question of one versus multiple disorders, Schaefer hoped that data on the symptom status of two such diverse groups of patients would at least clarify the issue of the appropriateness of a diagnosis of schizophrenia in each grouping. Forty-eight white male patients were located in a canvass of the files for the years 1943-55 (prior to the advent of tranquilizing drugs) of the Inpatient Psychiatric Service of the University of Minnesota Hospital. These patients were between the ages of 21 and 45, with a diagnosis of schizophrenia and had completed an MMPI within 10 days of hospital

admission. It is surprising that in the heartland of the MMPI only 48 patients over a 12-year span could be obtained who met these criteria for subject selection. Schaefer then applied the Phillips scale to the case histories that were available for these patients.

Of the 48 subjects, 32 were categorized as "poor premorbid" (on the basis of Phillips scores that were 18 and above) and 16 were categorized as "good premorbid" (on the basis of scores of 13 and below).

In addition to testing for the significance of differences between the groups on the 10 clinical scales of the MMPI, a comparison was made using Welsh's *A* and *R* factors (42), Barron's Ego Strength Scale (43), Schofield's Prediction of Change Scale (44, 45), and Jenkins' Prognostic Scale for Schizophrenia (46).

The profiles of the good premorbid and poor premorbid groups are presented in figure 1. There are strikingly few instances of marked differences between the groups. The "goods" differed significantly from the "poors" only on the L scale, whereas "poors" were significantly higher (0.05 level) than the "goods" on scales Pd (psychopathic deviate), Sc (schizophrenia), and K (validity scale). Since both Pd and Sc are weighted with K, the mean scale scores for each group without the K correction were computed. When this was done, both Pd and Sc failed to differentiate between the groups, suggesting that the obtained differences had been due primarily to the K weighting assigned to each scale rather than to differences that were basic to the contents of the specific scales. Using the method of non-K correction, the only reliable scale difference that was obtained was for the Ma scale (mania), with "goods" revealing significantly higher scores in comparison with "poors." Of the other scales introduced into the study, only Schofield's Prediction of Change Scale reflected reliable differences, with the "goods" scoring higher than the "poors." This find-

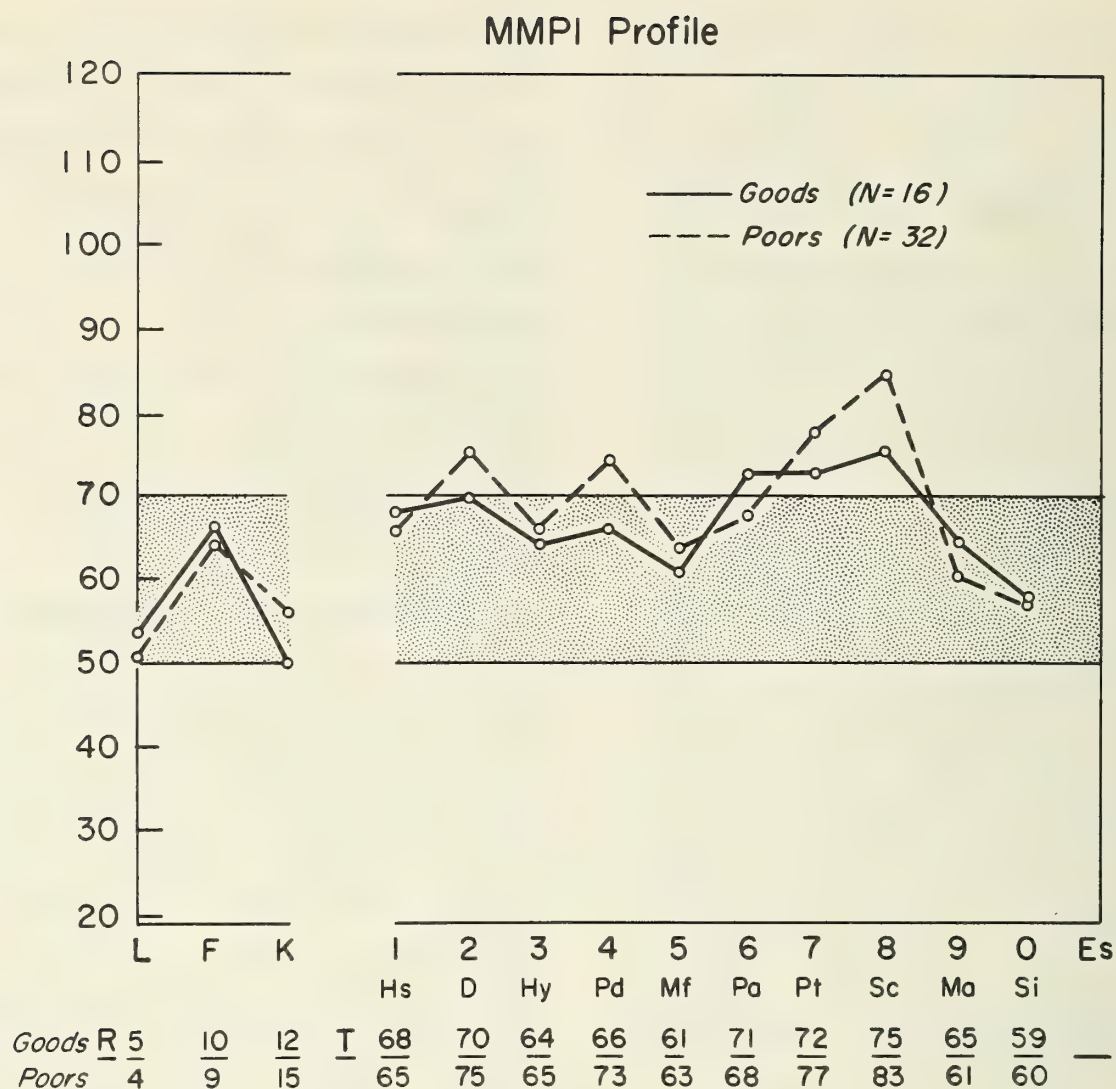


FIGURE 1.—Mean MMPI profiles of good and poor premorbid schizophrenic patients.
(From Schaefer)

ing is in keeping with the more favorable treatment potential of the former group.

In comparing the percentages of persons in each group who had scale scores ≥ 70 (the usual cutoff for more manifest pathology), no significant differences on any of the scales were found between the two groups of schizophrenic patients. In general, the mean clinical scale scores for the “poors” were higher than those of the “goods,” indicating a more severely disturbed sample of schizophrenic patients (a finding that is quite consistent with the characteristics of poor premorbid). However, there is a striking similarity in the configural patterns of the two groups.

Concurrent with, but independent of, Schaefer’s study, Church (47) also conducted a multiphasic study of good and poor premorbid male schizophrenic patients. Her groups, however, were far more chronic than were Schaefer’s. Here too,

however, the “poors” appeared decidedly more disturbed than the “goods,” sufficiently so for Church to question the applicability of a diagnosis of schizophrenia to all of her subjects, while calling for a replication of the study with more acute cases. Schaefer’s study, unconfounded by the effects of prolonged hospitalization or issues of diagnostic appropriateness, would suggest that these two groups of prognostically differentiated cases legitimately could be viewed as schizophrenic in terms of expressed symptomatology.

These data, however, do not provide a basis for the rejection of the view that there are multiple disorders at work in process-reactive schizophrenia. Although outcome for the two groups of patients may differ, symptom expression during the morbid (and acute) phase of the disorder, on which so much of diagnosis is based, does suggest the appropriateness of labeling so-

called process and reactive patients as manifestly schizophrenic.

2. *Process-Reactive: Dichotomy or continuum?*

Closely related to the first question is the issue of the identification of the process-reactive differentiation as a dichotomous typology or a continuum. The former viewpoint arose historically from the differentiations provided by Bleuler in terms of recovery/non-recovery. More recently, other investigators have urged that the process-reactive distinction should be viewed more appropriately as a continuum with the poles of the dimension represented by the classical process case at the one end and the reactive case at the other.

With regard to the dichotomous view, one can point to the distribution of scores of 343 schizophrenic patients rated by Wittman on the earlier Elgin Scale. Wittman's data are reproduced in figure 2 and attest to her view that "such bimodality indicates statistically that two definitely separated classes or types have been included in the sampling." Thus "the hypoth-

esis * * * that there are two types of schizophrenia differentiated as to prognosis, appears to be corroborated." (48).

By contrast, the view of a process-reactive continuum is derived from various sources of data. A number of studies report fairly substantial numbers of schizophrenic patients who fall intermediate between the two extremes. Kantor et al. in the paper that has been cited had two clinical judges rate the case histories of 108 schizophrenic patients on the basis of the 24 variables listed in table 1 (see p. 426). Judges then assigned each case to one of three categories: "process," "reactive," or "cannot say." Twenty-two of the 108 earned a "cannot say" rating. Since the records presumably were adequate enough for rating purposes, it would appear that ambiguity in classifying to either category existed in 20 percent of the cases. Additional evidence of ambiguity was provided by split ratings that occurred in 22 of the remaining 86 cases. Although some of this disparity may be accounted for by sparse case



FIGURE 2.—Frequency distribution of the prognosis scores for the 343 schizophrenic patients rated. (From Wittman)

history data, it is quite probable that many atypical cases were to be found within the 41 percent of the patients sampled who could not be consistently and accurately labelled.

A clearer example of diagnostic overlap is seen in King's (20) study of autonomic responsiveness in process-reactive cases. In this investigation, 90 patients were classified as either "process," "process-reactive," or "reactive," the intermediate category being designed "to represent schizophrenics with both types of characteristics to an almost equal degree." The characteristics which were rated were once again the 24 distinguishing attributes as suggested by Kantor et al. Apparently of the group of 90 cases, it was possible to secure clinical agreement on the assignment of 24 patients to each of the three categories. Thus approximately 21 percent of the patients selected could be reliably assigned to a combined process/reactive grouping, a figure in keeping with the original data provided by Kantor, Wallner, and Winder. Apparently the types of attributes provided by Kantor et al. are not distributed in the simple bimodal fashion suggested by Wittman (48).

Further support for the notion that a simple dichotomy will not suffice is apparent in the many laboratory studies that have been conducted with process-reactive or good premorbid-poor premorbid patients. The fact that extreme groups have usually been used for research purposes should predispose toward a bimodal distribution of data. Actually, an examination of the variability present within these groups clearly indicates the large degree of behavioral overlap that is characteristic of studies employing the two groups. Were evidence of such overlap to be restricted to one or two functions, one could argue that the response measures employed and the functions being tapped were irrelevant to the issue of a dichotomous typology. However, the range of studies evidencing overlap is so great and incorporates so many

behavioral measures that are presumed to be relevant to schizophrenic pathology, that it is difficult to continue to assert the existence of a simple process-reactive dichotomy. Simply to illustrate the range of contents in which such overlap between so-called process and reactive groups is clearly evident, one can point to studies of critical flicker frequency and spiral after-effect (17); genetic levels of perception on the Rorschach (19); childrearing attitudes of mothers and fathers as reported by schizophrenic and normal patients (49); size estimation of thematically meaningful pictures (50); verbal learning under reward and censure (51); abstract ability within the two groups (52); concept attainment on materials involving social approval and disapproval (53); and autonomic responsivity (54).

The list is as lengthy as the number of experimental publications that have appeared in the literature. One is restricted in tabulating such a listing solely by the depressing current tendencies of reporters of research to present data to the reader relevant to their measures of central tendency while failing to report their data regarding the dispersions of these measures. One is reminded of Kety's admonition that "recognition that any sample of schizophrenics is probably a heterogeneous one would seem to indicate the importance of analyzing data not only for mean values but also for significant deviations of individual values from group values."

The third form of evidence that could serve to establish a process-reactive continuum has meager experimental support. It has become fashionable in experimental psychopathology to use extreme groups when studying the effects of a given independent variable. In the case of process-reactive schizophrenia, it is not merely fashion that dictates this *modus operandi* but the far greater reality of the limited pool of schizophrenic subjects that is available for experimental purposes. The concept of a continuum, however, would neces-

sitate the use not only of extreme groups, but, at the minimum, the use of an intermediate group as well. This procedure was used by King in his study of autonomic responsivity that has been cited previously.

It will be recalled that King used the criteria provided by Kantor, Wallner, and Winder to select three groups of patients: process, process/reactive, and reactive. As a measure of autonomic responsiveness, the maximum fall in systolic blood pressure (MFBP) following the injection of mecholyl was used. Results indicated that the mean MFBP scores in ascending order of magnitude were as follows: process (lowest), process/reactive (intermediate), and reactive (highest). Further statistical tests revealed that the process group differed significantly from the process/reactive group with the latter, in turn, approaching a reliably lower value in comparison to the reactive schizophrenics. Unfortunately, efforts to replicate these findings without using an intermediate group were not only unsuccessful but, in one instance, produced diametrically opposing results (54, 55). This casts into doubt the reliability of King's results, but it cannot detract from the appropriateness of the use of an intermediate group of cases for experimental purposes.

Farina also has employed a categorization of subjects based upon both the extremes and an intermediate rating of poor and good premorbid schizophrenic patients, although his interest resided not in the patients themselves but rather in the patterns of dominance and conflict exhibited by the parents of good and poor premorbid schizophrenics. His comments on results obtained with parents of the intermediate group are interesting because they suggest a continuum of parental behaviors based upon variation in the premorbid adjustment level of their sons.

Inspection of the dominance indices of "good" and "poor" parents suggested that support for the hypothesis (of sex dominance of parents and

premorbid adjustment of the son) was more marginal in those cases where the sons' scores fell within an intermediate range (11-19) on the scale. Fathers of the "good" group appeared more clearly dominant over the mothers in those cases where the son had a very low premorbid score (toward the zero end of the scale) and seemed less dominant when the son's premorbid score approached 16. The data for the "poor" premorbid group seemed to show a more random variation. However, here too, the dominant parent (mother) seemed more clearly ascendant when the son's premorbid score was extreme (approaching 30) in comparison with the intermediate scale values. The net effect was to suggest that for the schizophrenic groups, father dominance increased with decreasing premorbid scores of the sons and mother dominance increased with increasing premorbid scores. In the middle ranges of premorbid scores (15 and 16) both parents appeared to be more nearly equated in dominance behavior" (56, p. 91).

Becker (24), in his research on genetic level Rorschach scores, employed a distribution of scores on the Elgin Scale and found a significant relationship between variation on the process-reactive continuum and developmental scores derived from the Rorschach as well as (in the case of males) the conceptual level of responses to the Benjamin Proverbs Test. More process-like schizophrenics obtained lower genetic-level Rorschach scores; more reactive-like schizophrenics secured developmentally higher scores. Results such as these add support to the likelihood that the process-reactive dimension represents a continuum of severity of the schizophrenic disorder rather than a simple dichotomous entity. Most results obtained in studies (e.g., Slotnick, Solomon and Zlotowski, etc.) employing correlational methods in which the variables represented are the patient's level of process-reactive status on the one hand and variations in some behavioral measure on the other reflect assumptions that a continuum rather than a typology is

most appropriate for characterizing the severity of the schizophrenic disorder.

3. *Process-Reactive: Organic versus psychogenic etiology?*

A persistent and unyielding conceptualization of the process-reactive continuum involves the assumption that the primary etiological factor in the former is biological/genetic whereas the antecedent for the latter is primarily psychogenic. I have already cited reasons for the tenacity with which this view is held. What are its origins? And how viable is the belief in the light of available evidence?

The source of this orientation is clearly evident in Bleuler's classification of mental diseases. Four decades ago in the English edition of his "Textbook of Psychiatry" Bleuler wrote:

One class of the psychoses shows itself as a morbid reaction to an affective experience, as a prison psychosis to a confinement, and an hysterical twilight state to a jilting on the part of the beloved (reactive psychoses, situation psychoses). In the other class there is a morbid process in the brain, that conditions the psychosis (process psychosis, progressive psychosis). But no division can be based on these classes because the two symptomatologies intermingle (3, p. 173).

Certain inherent biases in this paragraph are suggested. "Process" implies an ongoing, progressive, organically-determined development over time; "reactive" implies a more immediate response to some current stimulus or influencing condition. The etiologic assumptions are also clear as is the dichotomous classificatory schema involved. However, the wisdom of Bleuler is most clearly seen in his cautionary note that an element of interaction is present in the etiology of both process and reactive conditions since organic and psychogenic elements intermingle. Nevertheless, for reasons of the prevailing biases that surround the terms "process" and "reactive," other investigators have seen fit to select the more

neutral "good" and "poor" premorbid terminology since these are operationally specific to the scale used for rating patients. In commenting upon this decision, Garmezy and Rodnick wrote:

To the reader the separation of patients into "good" and "poor" premorbid categories may suggest that we too have a dichotomous conception of the disorder which is coterminous with the process-reactive differentiation. This is not so. The separation of patients into "goods" and "poors" is not founded on a theoretic belief involving differentiated etiologies but is simply a first entry into the problem of reducing variance through the use of a scaling instrument with marked limitations. Although "goods" and "poors" do share some attributes which have been assigned to reactive and process cases, our own position is theoretically neutral and makes no assumptions involving an exogenous-endogenous position. Indeed the research of our group strongly suggests that premorbid adjustment can be thought of as a continuous rather than a dichotomous distribution. In several recent studies the use of the scale in a trichotomous fashion provides further differentiations within the groups. For experimental convenience, however, we have taken to using the extremes of the distribution — an arbitrary but useful procedure. Therefore, our own view of "goods" and "poors" is based upon utilitarian considerations. The critical empirical problem which remains unanswered is that of determining the nature of the range of differences in performance between the groups in various tasks and the exploration of personality variables and life history factors which are correlated with such differences (40, p. 454).

Evidence in support of a differentiated etiology is based upon two classes of data: the first relates to observed similarities between cases of clearly defined organic pathology and chronic (presumably process) schizophrenia; the second relates to genetic determinants in process versus reactive schizophrenia.

Typical of the first group of studies comparing organic and chronic schizophrenics are those related to conceptual performance. Tutko and Spence (58), in a carefully designed study, compared the performance of normal, brain-damaged, and process and reactive patients (on the basis of Phillips' scores) on the Goldstein-Scheerer Sorting Test. Their hypothesis was that the similarity between organic and process cases would be greater than that obtained for the process and reactive groups. Of greater importance was the formulation advanced by these investigators that the types of errors made by the groups would differ.⁸ For process and organic patients, Tutko and Spence predicted that the errors made would be primarily restrictive ones ("an inability to suggest similarities among sets of objects"); for reactives, errors would be primarily of the expansive type ("inappropriately broad or idiosyncratic"). Typical of restrictive errors were those categorized as "no response," "split-narrow" (the definitions embrace only a subgroup of items), "chaining or object naming," and "concrete fabulization" (objects are merely mentioned in the context of a narrative). There were two types of expansive errors: abstract fabulization (an extended narrative related or unrelated to the objects) and syncretistic and symbolic responses (overinclusive loosely associative boundaries).

The findings of the study were essentially these: (1) Normals gave the greatest number of abstract responses but did not differ reliably from the other three groups; (2) reactives' errors were primarily expansive ones; (3) the errors of the process and brain-damaged groups essentially were restrictive in nature; (4) the reactive group was most prone to give additional responses followed by the normals, process, and brain injured.

⁸ I do not mean to imply that Tutko and Spence have been the only investigators to study conceptual deficit in terms of specific response attributes. The research of Cameron (59), Epstein (60), McGaughran and Moran (61), Payne (62), and others clearly belie this.

Summarizing their study, the authors note:

In most of their Inadequate responses, the brain-injured subjects erred in the direction of being unable to specify a common property that characterized all of the objects. That is, they failed to suggest any basis of commonality, specified a property of only some of the objects in a group, or merely described them in some fashion. However inadequate their responses often were, their approach to the task might at least be characterized as appropriate in that they tended to stay within the limits of the task as specified by the experimenter's instructions and confined themselves in their responses to the fairly immediate properties of the objects.

The performance of the Process schizophrenics was similar to the brain injured in that the predominant types of error made by these subjects were of the same kind. However, their inability to specify common properties of the objects was less marked than the brain injured. That is, they gave significantly more Adequate responses and more often suggested multiple bases for the objects belonging together. Further, for those responses that were considered Adequate, the Process schizophrenics gave more responses that were categorized as Abstract.

Like the Process subjects, the Reactive schizophrenics gave significantly more Adequate responses than the brain injured, and within this category, proportionately more Abstract responses. In both of these measures, however, the two schizophrenic groups did not differ significantly and thus may be said to have been equivalent in overall goodness of performance. There were marked differences between the groups, however, in the kind of Inadequate responses given. The Reactives might be described as erring in the opposite direction from the Process and brain injured. The Reactives were typically able to suggest some basis for the objects belonging together (response failures were rare). Their responses, however, were frequently overinclusive in the properties they specified, loosely associated with the specific objects, or so bizarre and

idiosyncratic as to have no discernible connection with the objects. The willingness of the Reactives to offer a response, however inappropriate, is further demonstrated by the finding that they gave significantly more multiple responses than any other group, including the normals. This frequently had the effect of saying too much, since within a single answer, it was not unusual to find a high level, abstract response among responses of poorer quality.

The overall performance of the normal control group was, of course, superior to all of the other groups; i.e., they gave a significantly greater number of Adequate responses and a higher percentage of Abstract responses. Examination of the Inadequate responses of this group revealed less of an imbalance between Restrictive and Expansive errors than in any of the other groups. The majority of their errors, however, were of the Expansive type and in this sense, they were more similar to the Reactive schizophrenics than to the other groups (58, pp. 392-393).

Although the authors suggest that their findings may "support the hypothesis that some type of brain pathology may be responsible for the symptoms exhibited by process schizophrenics, as opposed to etiological factors of a more functional nature in reactive schizophrenics," they are appropriately cautious in suggesting that different etiological considerations may produce similar symptom pictures.

This observation, of course, is one that was made by Jaspers in his monumental work on general psychopathology:

That the *outcome is the same* is no proof that the *disease is the same*. The most diverse organic brain diseases show the same outcome in the same type of demented state (63, p. 569).

I have quoted Tutko and Spence at length in order to give the reader the opportunity to consider their findings in some detail. It is clear from their description that process and organic patients differ in certain significant ways. The organic patient apparently attempts to grapple with

the sorting problems but is *cognitively* unable to do so successfully. Furthermore, this incapacity is accentuated by the fact that he is apparently able to maintain the task set. By contrast, the process patient is conceptually more adequate, shows an ability to produce more adequate and abstract responses than his organic counterpart, but his performance appears to reflect more a *motivational* than a cognitive insufficiency. The restrictive errors to which Tutko and Spence make reference are the more probable kinds of errors that an inactive and withdrawn patient is likely to make. A low activity output combined with an inability to maintain a task set are attributes that are more likely to generate non-responding, or attention only to subsets of stimulus materials, or object naming, or simple narrations. To this extent, one can perceive in the conceptual habits of patients a continuity with those behaviors of the premorbid period that are so characteristically process-oriented. By contrast, the most striking attribute of the performance of reactives is their "coping behavior." Despite the destructive impingement of their psychosis, they continue to search for solutions to the problems, however inappropriate these efforts may be. If they speak too much, they do at least try unremittingly to attain solutions and such efforts to grapple with problems also suggest a continuity with their more adaptive patterns of the past.

This point perhaps can be made clearer by referring to Payne's systematic studies of overinclusiveness (62). Various tests of conceptual performance (including the Goldstein-Scheerer Sorting Test, Payne's Object Classification Test, and Benjamin's Proverb Test) suggest that overinclusiveness is characteristic of acute schizophrenics whereas chronic schizophrenic patients do not show such behavior. What are the criteria for overinclusiveness in these tests? For the Goldstein-Scheerer Test, overinclusiveness is measured by the number of objects grouped together in the "handing over" part of the test; for the Object Classi-

fication Test (which provides for 10 correct sortings), the measure is in terms of the variety of additional unusual sorts given; for the Proverb Test the response measures are in terms of both talking time and the number of words spoken. These responses too are activity measures and thus it is not surprising to find acute patients manifesting overinclusive behavior whereas chronic patients do not. In comparable fashion, Epstein's overinclusiveness score was a measure of frequency of responding with items that were presumably an essential component of the stimulus concept.

In making this point, it is not my intention to negate the significance of overinclusiveness when applied to schizophrenic behavior; but I do suggest that it is important to consider the nature of a response measure and the diverse processes that may underlie it when evaluating seemingly comparable performances by process and organic patients. A combination of inability to maintain a task set, the intrusion of irrelevant non-task oriented stimulus events and a depressed behavioral output can produce a picture of conceptual deficit in schizophrenia alike in certain structural respects to that produced by organics, but with strikingly different process components underlying the behavior.

However, other demurrers to the organic-process parallelism can be offered. Some investigators have reported negative evidence regarding the performance of process and brain-damaged patients on tasks that appear to be relevant to a concept of organicity (17). Others find that obtained differences may be explicable on factors related to variables such as length of hospitalization rather than process-reactive status (64). This last point brings to mind the observations of Kety (65) and Horwitt (66), that obtained physiological differences between schizophrenics and non-schizophrenics often reflect consequences, rather than antecedents, of psychopathological processes in schizophrenia.

In Horwitt's words:

. . . It is earnestly hoped that investigators, impelled to study the biology of schizophrenia or of other mental disorders, will attempt to control the variables mentioned so that we may better distinguish between the causes of schizophrenia and its effects. Admittedly such controls are expensive and difficult to administer, but they are worthy of incorporation into any research program where man is the experimental subject. Much has been said about the faults of psychiatrists who do not make sufficient use of the laboratory concepts of cause and effect in evaluating mental disease. Conversely, the biologist should not be so naive in the interpretation of his data that he loses cognizance of the fact that schizophrenia is not a simple entity, and that he, too, must beware of the trap of confusing cause and effect (66, p. 430).

Kety's concern is equally strong:

The physiological and biochemical changes which are secondary to the psychological and behavioral state of the patient are of interest in themselves, and understanding of them contributes to total understanding of the schizophrenic process; it is important, however, not to attribute to them a primary or etiological role (65, p. 1159).

Such admonitions to the biologically-minded have equivalent significance for those who maintain a more psychological posture toward schizophrenia.

Genetics

A more telling point that suggests the potential validity of a process/organic-reactive/psychogenic hypothesis is contained in the literature of the genetics of schizophrenia. But here too the literature is, at best, sparse and equivocal. The most relevant data are to be found in Rosenthal's (67) superb analysis of that genetics literature.

Using Slater's data on identical twins (68) as the base for this analysis, Rosenthal

compared twin pairs that were concordant and other pairs that were discordant with respect to schizophrenia as a test of his hypothesis that "typical or process schizophrenia is found more frequently among concordant pairs, atypical or reactive schizophrenia occurring more often among discordant pairs of one-egg twins." Slater's case histories were rated against Phillips Scale criteria to ascertain each twin's premorbid social and sexual adjustment. Unfortunately, all of the 11 sets of twins produced Phillips scores that fell toward the poor premorbid end of the scale. For the concordant pairs the mean Phillips Scale for the index case was 23.2 with the mean for the co-twins 24.4; for the discordant pairs the means for the index cases and co-twins were 21.3 and 16.0 respectively. Thus the absence of reactive schizophrenics in the sample does not allow for a determinate test of Rosenthal's hypothesis. In support of his contention is the observation that four of the six sets of twins within the discordant group who had escaped the illness had the better premorbid social and sexual history. However, only one twin of the six sets had a premorbid score that would unequivocally place him in a good premorbid category.

A further analysis of the data revealed that severity of illness was linked to concordance. Using age of first admission as an index of onset, Rosenthal compared his concordant and discordant pairs and found the mean age of onset in the former to be 22.9 years versus 37.7 years for the latter group. Furthermore, *there was virtually no overlap between the two groups*. In terms of outcome data, "poors" tended to show a more deteriorative course than did the "goods." Since onset and outcome are more determinate indices of process-reactive status, Rosenthal's data appear to suggest the presence of a genetic bias in the process group and a non-genetic component among reactives. The greater prevalence of schizoid-schizophrenic illnesses among family

members of "poors" in contrast to the "good" premorbid group would appear to lend further substance to the data. Two additional points, however, should be noted. First, Rosenthal's data on female schizophrenics failed to support the triad of hypotheses of greater severity of illness as well as earlier age of onset and more baleful outcome in the concordant twins. Second, preliminary data gathered by Gottesman and Shields,⁹ who are currently conducting a twin study of 57 pairs of twins drawn from 16 years of consecutive admissions seen at Maudsley Hospital during the period 1948-64 suggest that the concordant-discordant pairings are not readily explicable on the basis of Phillips premorbid scores. Other data, however, are supportive. Gottesman writes:

We found no strong connection between age at first hospitalization and MZ concordance. Before age 25, 7/13 were concordant, with 7/15 concordant after age 25, counting probands rather than pairs. By omitting probands with an onset of schizophrenia at age 35 or later, concordance rates are 14/25 (56%) for MZ and 3/18 (17%) for DZ. Among our 10 concordant MZ pairs 5 had onsets after age 20. Among our 14 discordant MZ pairs 9 had onsets after age 20.

Probably our most interesting preliminary findings that tie up with your notion of a process-reactive dimension are those relating severity to concordance. Dividing the MZ and DZ probands into those hospitalized for under and over one year, we find concordance rates of 20% and 67% in the MZs and 11% and 12% in the DZs. Using under and over two years, we find 27% and 77% in the MZ and 10% and 15% in the DZ.

With work situation at the time of follow-up as another indicator of severity we obtain a confirming view of the relationship with concordance. For MZ probands working and out of hospital for more than six months when last known, the concordance rate is only 17%. For all other outcomes the

⁹ Personal communication from Gottesman and Shields.

MZ concordance was 75%. For DZ probands with a good outcome the concordance rate was 0% and for all other outcomes concordance rate was 22%.

Relationship between hospitalization and concordance did not appear to be related to current age.

Gottesman's current analyses are only tentative and an answer to the issue of genetic variation and biological heterogeneity in schizophrenia must await a more definitive treatment of his data.

For the present we are left with Rosenthal's conclusion that—

* * * the combined findings of discordance in monozygotic twins and the virtual absence of schizophrenia among members of their families are strongly suggestive that a schizophrenic gene is not the responsible agent in the illness of these twin index cases * * *

On the other hand, the combined findings of concordance in the monozygotic twins plus the fact that history of probable schizophrenic illness occurs in approximately 60 percent of their families is strongly suggestive of an hereditary determinant (67, p. 8).

Rosenthal's findings are suggestive; hopefully Gottesman's data may shed still more light on this issue. But the question of biological heterogeneity in schizophrenia is obviously extremely complex, and among identical twins variations in premorbid adequacy could well be related to variations in environmental stress. Recent publication of the volume devoted to the Genain quadruplets (69) can well serve as a case in point.

Most case histories that are available in the genetic studies of schizophrenia are so sparse and so unyielding of case dynamics that factors in the premorbid history are generally limited (and often non-productively) to more demographic forms of data. It seems reasonably conservative to assert that the magnificent case study of the Genain quadruplets provided by Rosenthal, Raphling, and Quinn in that volume and

by Wynne et al. (70) in a separate publication stand alone in the literature of the genetics of schizophrenia. Without such a history it would not have been possible to perceive how the different role relationships established during the earliest years of the quadruplets' lives set in motion varying patterns of adaptation that were to eventuate in a process-type of schizophrenia in three of the four quadruplets and to produce a more reactive form of the disorder in the fourth.

As Rosenthal points out in his effort to erect a conceptual framework for the study, theories that emphasize a diathesis-stress interaction are coming to prominence in research in schizophrenia. Such theories (71) assert that a predisposition to schizophrenia is inherited but that the disorder more typically is potentiated by a variety of stressors—more often psychological in nature. A review of the life histories of the Genain quadruplets not only implicates heredity but points to differential environmental experiences as a potential component in the subsequent development of a more process-like syndrome of schizophrenic symptomatology. But studies of varying psychological antecedents in process and reactive schizophrenia are of fairly recent origin. Since many of these studies suggest that variations in the role patterns within the family of schizophrenic patients not only exist but can be differentiated along a process/reactive dimension, a brief review of the literature is in order.

Before doing so, I feel impelled, in closing this section, to quote Ralph Gerard's words of wisdom:

Avoid the disaster and confusion that results from the careless admixture of different levels of discourse. Mind does not act on matter, nor matter on mind. There are only an antecedent mind-body state and a consequent mind-body state, whether mental or physical aspect chances to present more acutely * * * This can all be given point at a favorable level of argument, as to the cause or genesis of psychoses.

The constitutionalists and the organicists and environmentalists and mentalists too often are quarreling with each other as to which of them has the cause. Now it is obviously useful to find out that schizophrenics have abnormal capillaries in their fingers, that they had abnormal experiences in their childhood, and that they have abnormal individuals as parents and sibs; but one does not exclude the other and no one of them can possibly be the whole story (72, p. 3).

4. *Studies suggestive of differential patterns of socialization in process-reactive schizophrenia*

There is a growing, if inconclusive, body of evidence to suggest that variations in premorbid social and sexual adequacy may be related to variations in family organization and socialization practices. Such data, some of a direct nature, others highly inferential, may have implications for variations in early socialization patterns that characterize process and reactive schizophrenics. Several of these studies have been reviewed elsewhere (40) and, therefore, I shall merely catalog some of these recent findings.

- Poor premorbid patients, when responding to a child-rearing attitude scale as they believed their mothers and fathers would have responded during a period of time when they (the patients) were growing up, tend to ascribe more pathologic child-rearing attitudes to both parents. Intrafamilial comparisons based upon such items suggest a pattern of heightened maternal dominance in families of "poors" and more characteristic patterns of paternal dominance in the case of "good" premorbid schizophrenic patients. Thus the traditional pattern of powerful mother/ineffectual father that is so pervasive in the familial literature of schizophrenia appears to be more typically a "poor" premorbid pattern (49).

- Actual parental interactions in a situational test procedure confirm the above findings: (1) For the "poor" pre-

morbid group the mother reveals marked dominance with father playing a more submissive role; a characteristic pattern of parental conflict and discord predominates; (2) By contrast, the "good" premorbid parental groups tend toward father ascendancy with mother showing greater submissiveness; a lesser pattern of discord is in evidence; (3) normal control parents share authority patterns and reveal little overt indication of conflict (56).

- "Poor" premorbid patients reveal greater deficits in visual discrimination and concept formation if the stimulus contents involve maternal censuring cues; there is a lesser tendency for "goods" to be affected by pictorial cues depicting father as the censuring figure (73, 74, 75).

- "Poor" premorbid patients, in terms of response to TAT cards, show greater "anxiety-related imagery" in response to mother and a lesser response to father; "goods" show a reverse pattern. "Poors" show greater avoidance to both parental figures (76).

- Parents of "poor" premorbid schizophrenic patients reveal more immature defensive behavior on the Rorschach Test than do parents of "good" premorbid patients; parents of neurotic controls fall intermediate between the two schizophrenic groups with the differences between "poors" and neurotics tending to approach statistical reliability (77).

- "Good" premorbid schizophrenic patients show greater interference in performance on the Digit Symbol Test of the WAIS following exposure to a tape recording of a father censuring a son than do "poors." By comparison, "poors" show the greater interference following maternal censure. Intragroup comparisons indicate that "goods" are more affected by paternal than maternal censure whereas "poors" show a reverse pattern. "Goods" tend to be more disturbed by father-son conversations, particularly under conditions in which the parent-child interaction is characterized by censure (78).

In citing these results, I must once again assert that I am not attempting to negate the importance of biological heterogeneity in schizophrenia. For one thing, data such as those cited are fraught with shortcomings that enable only the most adventuresome investigators to assume that they unequivocally indicate the existence of distorted family pathology early in the life of the preschizophrenic child (79). Nevertheless, they do allow for some interesting speculations (49). Marked maternal dominance tends to produce male children who are docile, submissive, excessively obedient, and estranged from their peers (80). Such behaviors are not unlike those manifested by "poor" premorbid patients. Heightened conflict between parents of "poors" may further accentuate the social withdrawal of those children who later develop a more malignant form of schizophrenia.

On the other hand, the "good" premorbid patient may often have a father of more masculine stature who at least provides the patient (as a child) with a model of masculine behavior although their strong tendencies to punish or threaten may accentuate the predisposition to anxiety that is characteristic of the good premorbid schizophrenic patient. Such an earlier pattern could also help to explain why the good premorbid patient often tends to develop a psychosis under intensive threats to his self-esteem and masculine identity. In any event, it is difficult to assert unequivocally that an organic substrate underlies a process schizophrenia and a psychological substrate the reactive type when there is some reasonable evidence suggesting that different psychological antecedents may also be at work in the two forms of schizophrenia.

Three decades ago, in introducing his survey of existent research in dementia praecox, Nolan D. C. Lewis was impelled to write:

* * * There still is too great a tendency on the part of the workers to consider mental disorder either as due to

influences penetrating from without, or due to something dissolving or disintegrating within the individual. Even those who express themselves so as to include both general sets of factors, usually formulate problems which are pointed toward either the "psychogenic" or the "organic" concepts in psychiatry—to use a current terminology (81, p. 1).

The terminological distinction continues to exist often, unfortunately, accompanied by the same biases of an earlier period.

Behavioral Heterogeneity and the Process-Reactive Continuum

Were the distinctions produced by the use of the Elgin and Phillips' Scales to be restricted solely to prognosis, the scales would have a limited import. The fact, however, that such distinctions are now coming to prominence either as independent or control variables in a wide-ranging series of research investigations further heightens the importance that must be accorded the continuum. To catalog all references to the construct validity of these scales would extend this chapter far beyond its present limits. The reader can find testimony to the utility of the distinction in a number of publications including those of Higgins (38), Herron (39), Garmezy and Rodnick (40), Rodnick and Garmezy (30), and Silverman (82), to name but a few.

Study upon study has reported reliable differences between process-type and reactive-type cases in such diverse areas as learning (51), cognition (58), language and verbal behavior (83, 84), perception (85), motivation (86), judgment (87), physiological responsivity (88), child-rearing attributes of parents (49), compliance and avoidance reactions (89, 91), adaptation to social rewards and punishments (89), parental relationships (90), etc.

Merely as illustrative, but unequivocally effective as an example of the utility of the Phillips Scale for reducing behavior variability in schizophrenia, is the research on

size estimation conducted by Harris (50). Harris' interest lay in relating the symbolic content of pictorial stimuli to distortions in the size estimates accorded these stimuli by schizophrenic subjects. To test his hypothesis of a tendency by patients to overestimate, Harris constructed scenes that depicted a mother and a son in several different interactions — overprotective, rejecting, and accepting—together with a more neutral tree-bush scene and a simple geometric figure (a square). The patients first viewed the scenes and were then asked to judge from memory the size of the pictures they had been shown.

The data for the total schizophrenic and normal groups showed such marked overlap that statistical tests are not required to reveal the sterile quality of the findings. Figure 3 provides the picture of nonsignificance. When, however, the total pool of schizophrenic patients is separated into subgroups of "good" premorbid and "poor" premorbid patients, the differences are dramatically accentuated. As can be seen from figure 4, "poor" premorbid patients show the predicted overestimation of mother-son pictures whereas "good" premorbid patients tend

toward underestimation; by contrast, normal subjects are clearly revealed as the more objective perceivers of the size of previously seen pictures. Commenting on the productive returns realized by Harris' use of the Phillips scale, Garmezy and Rodnick wrote:

What if Harris had failed to separate his patient group? In that event, his results would have proved disappointingly ambiguous * * * Combining the data of good and poor premorbid Ss results in mean values which approximate the objective reality achieved by normals in estimating the size of a standard stimulus; but this is clearly an artifact produced by summing the widely disparate and bidirectional behaviors of the two subgroups of patients. Had Harris, then, viewed the schizophrenic group as a totality, he would have committed a type 1 error (the rejection of a hypothesis that is true). To return once more to the problem of variability, it is of importance to note that the combined schizophrenic group showed such marked variability in performance that a comparison of the variances of the experimental and control groups by F tests for each scene necessitated, in

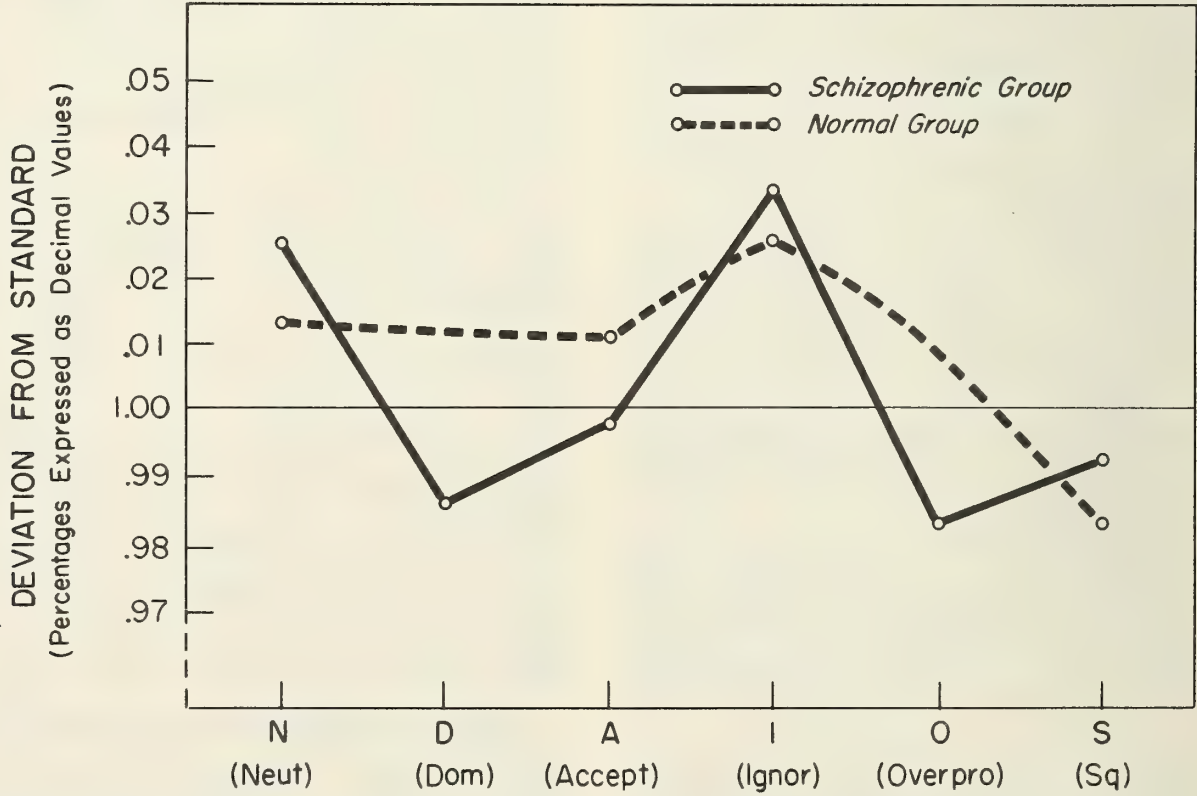


FIGURE 3.—Mean size estimates of each of the six scenes for total schizophrenic and normal groups. (From Harris)

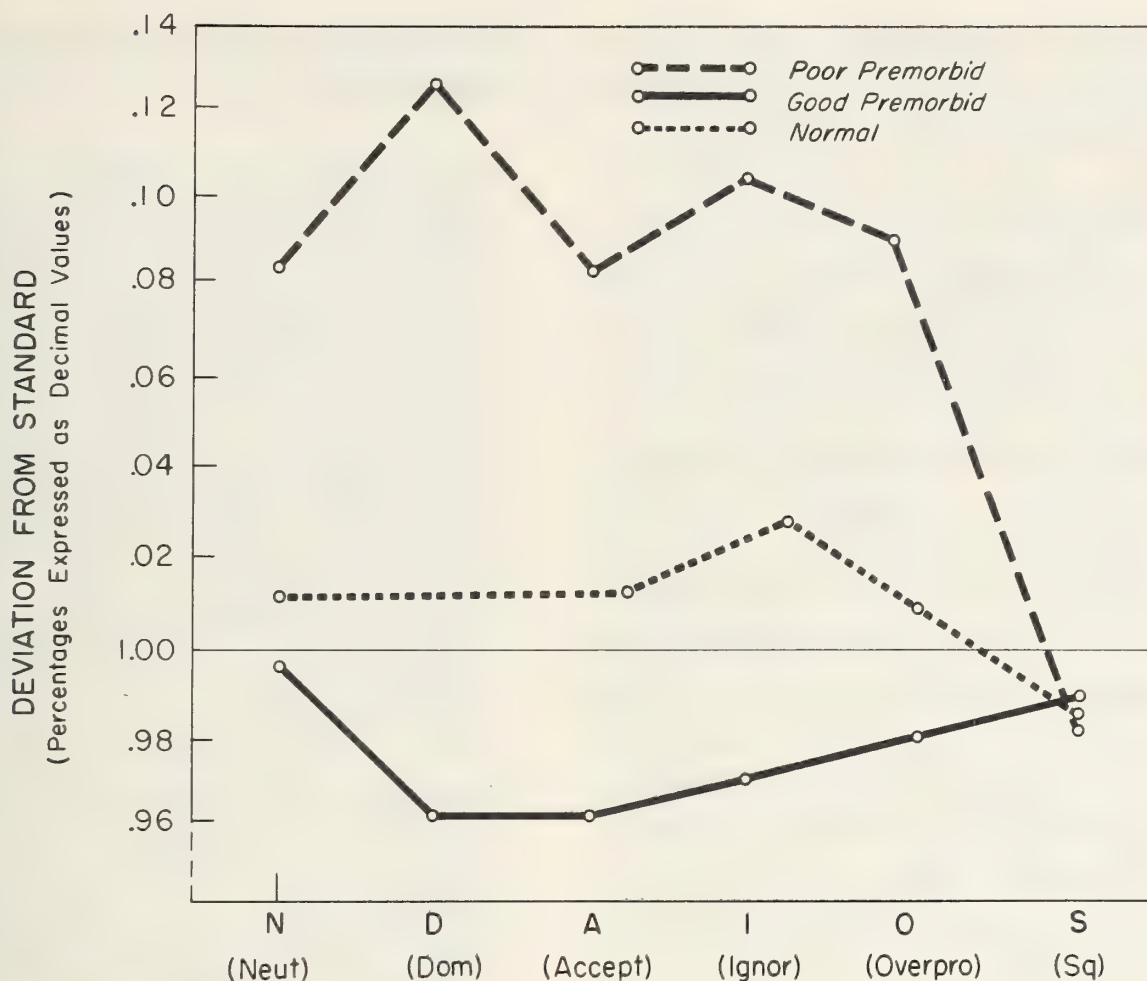


FIGURE 4.—Mean size estimates of each of the six scenes for good and poor premorbid schizophrenic and normal groups.
(From Harris)

most instances, rejecting an assumption of homogeneity of variance—a condition which did not obtain when the good and poor premorbid separation was employed (40, pp. 459–460).

Of greater significance than the point cited above are the implications for theory that can be adduced from data supplied by more homogeneous subgroups of schizophrenic subjects. Again the Harris experiment is a case in point. Recently, Silverman (92a) has advanced a cognitive control theory of attention which permits an effective integration of Harris' research with other seemingly unrelated studies in schizophrenia. Silverman relates the tendencies toward overestimation to Piaget's formulations of developmental factors in relation to attention. "Centration"—the tendency to fixate upon a stimulus—heightens overestimation of the stimulus. In the young child such centration effects are the result of a tendency to anchor upon

dominant objects in the stimulus field—a tendency that is reduced as the child grows older and learns to shift his attention toward and away from the center of a stimulus field. Individuals differ in the extent to which they scan visual fields; extensive scanners show little over- or underestimation, but limited scanners tend to overestimate the stimulus objects to which they attend. Such response dispositions when applied to schizophrenics, Silverman asserts, can be related to symptom, premorbid history, and acuteness of the disorder. "Poor" premorbids are minimal scanners, "good" premorbids show more extensive scanning behavior. Such developmentally earlier scanning behavior is consistent with differences in genetic levels of perceptual development between process and reactive cases as suggested by the Rorschach developmental scoring studies of Becker (24) and Fine and Zimet (19, 92).

Of comparable importance are the defensive aspects of these two variants in scanning behavior (93). Minimal scanners may be more avoidance-oriented in the face of stressful stimulation, turning away from threatening aspects of their environment and reducing their field of perceptual and cognitive awareness. Such behaviors are clearly consonant with the more acutely ill "poor" premorbid patient's hypersensitivity to potentially noxious stimuli and to his prepotent habits of defensive avoidance under threat (89). The excessive scanner Silverman relates to the hyperalert individual who searches for cues that arouse his anxiety and then seeks to cope with such arousal by the production of cognitions and ideations that require a reinterpretation of potentially noxious stimulus events—behavior that is more characteristic of a paranoid schizophrenic.

To return to the Harris experiment, I would merely assert that it is not a unique instance of the success achieved by use of the Phillips scale in reducing behavior variability while clarifying varying patterns of response in the two premorbid groups. But is the Phillips Scale (and by implication the Elgin Scale as well) unique in its capacity to generate such behavioral homogeneity? Can a differentiation based upon a more traditional Kraepelinian system of classification produce equivalent effects? This important question warrants discussion.

Diagnostic Subtype, the Process-Reactive Continuum and the Reduction of Behavior Variability

Recognition of the prognostic variations among the subtypes of schizophrenia extends back to Kahlbaum and remains a continuing concern of psychiatry (94).

The search for application of the Kraepelinian classification schema to research in schizophrenia can be seen in the following statement from the most recent edition

of Henderson and Gillespie's "Textbook of Psychiatry":

Research into the many problems posed by schizophrenia has been frustrated by doubts and uncertainties about the heterogeneity of the clinical material being studied. Possibly the majority of observers now think that schizophrenia is not a clinical entity but a group of syndromes, of different aetiologies though with similarities in their clinical features; yet research has usually proceeded on the assumption that a biological entity was being investigated. The varieties of schizophrenia shortly to be described are clinical groupings of symptoms, they have no known pathological basis; but it might well now be more profitable to select for biochemical research these clinical subdivisions rather than schizophrenia as a whole, and particularly the most typical cases of hebephrenia and katatonia (95, p. 253).

This suggestion had been long realized in the early interdisciplinary studies of schizophrenia that had been conducted at Worcester State Hospital during the 30's and 40's. Under the direction of Hoskins, Shakow, Malamud, and others, a series of physiological and psychological studies were conducted primarily on long-term chronic cases (96). The psychological investigations, in Shakow's words, "ranged from the patellar-tendon-reflex latent time at one extreme to group behavior involving competitive and cooperative activity at the other * * *" To review these studies would take us far afield, but it is clear that in many areas the use of clinical subtypes produced marked differences in adequacy of performance. Thus, paranoid patients revealed superior performance on a test of steadiness (97); catatonics showed the greatest improvement over mean reaction time in successive experimental periods (98); on a pursuit learning task the order of achievement on initial performance (from highest to lowest) was: paranoid, unclassified, hebephrenic, indeterminate, and catatonic (99).

In a recent review of part of the Worcester program of research Shakow was prompted to write:

Both interindividual and intraindividual *variability* are a major source of difficulty in research on schizophrenia. In our psychological studies, groups of schizophrenic patients have quite consistently given coefficients of variation three times that of normal subjects, and individual patients two times that of normal ones. In studies of physiological functions, these have varied from one and one-half to more than twice that of normal subjects. Of the many sources of variation I shall only concern myself here with the two important ones of nosology and attitude.

Taken in its broadest sense, *nosology*, has many facets. Initially there is the fundamental problem of the diagnosis of schizophrenia and its subtypes. Aware of this problem, we have made every effort, especially in our Worcester studies, to obtain reliable diagnoses of patients. In addition to certain definite exclusion criteria, we had specified standards for the general diagnosis of schizophrenia, as well as for subtype classification. Whenever there was a question about the schizophrenia diagnosis, the patient was not used in the research. If the patient did not clearly meet the criteria established for one of the four subtypes, the category of mixed, unclassified, or indeterminate was used.

We have a tendency these days to be condescending about subtype classification. In fact, there are even "nihilists" among us who would do away with all diagnostic categories, the less enthusiastic of these calling for at least the casting overboard of the subtypes. If diagnoses are to be made on the relatively careless bases so prevalent in many centers, then I would go along with this point of view. If, however, they are based upon carefully worked out and tested criteria, then they deserve considerable respect. I say this despite my original and continuing "dynamic" bias. For the clear-cut syndromes of behavior, for which such labels are referents, have definite value for research purposes, even carrying consistent dynamic and "style" im-

plications for those interested. I am not claiming that there is not much room for improvement. Perhaps it will come through the factorial techniques, which, after some thought, we decided not to use at Worcester. What I do hold is that such perceptive clinicians as Kraepelin and Bleuler saw things that we might also see if we look carefully (100, pp. 275-276).

The elaborate use of subtypes, characteristic of the Worcester research, has not been followed in most psychological studies of schizophrenia. A sampling of the literature does indicate that a substantial number of studies have employed a paranoid/non-paranoid dichotomy without attending to a more thorough division by diagnostic subtypes. Silverman (82, 92a, 93) has cited and described many of these studies (101, 102, 103, 104, 105, 106 107, 108, 109, 110).

A number of factors that extend beyond theoretical considerations may account for the active interest in the paranoid/non-paranoid differentiation. For one thing, paranoid patients are more readily identified psychiatrically and thus constitute at least a partially homogeneous group. For another, they bulk large as the subjects for experimental investigations since their relative intactness, frequent cooperativeness and ideational sufficiencies make them ideal choices as participants in research. Third, they are sufficient in number in many large hospitals to provide an adequate *N* for a designated experimental patient group. This is not to deny the importance of such a distinction, but rather to indicate factors that have restricted the use of the entire range of the Kraepelinian classificatory schema in schizophrenia. Still another factor, of course, has been the notorious unreliability of psychiatric diagnosis. In many hospitals the very careless bases of classification to which Shakow alludes reflect the more typical diagnostic attitudes. Only in research-oriented psychiatric centers where either the clinical staff shares a dedication to the research enterprise or the research staff carries on its own program

of vigilant psychiatric diagnosis is it possible to approximate the diagnostic exactitude and acumen that was so characteristic at Worcester.

Does this mean, however, that given the same meticulous attention to diagnosis a comparable research payoff to the one achieved in the Worcester studies would result? I think not, and the reasons for this view are contained both within the attributes of the subject samples that were used in those earlier studies together with a particular methodological procedure that was uniformly followed.

With regard to the subject attributes, Shakow writes:

The patients we generally studied can be described as chronic. They had a mean age of approximately 30, a mean schooling of 9 to 10 years, and a mean hospitalization age of approximately 7 years. (Hospitalization age is defined as the time elapsed since first hospitalization for mental disorder (100, p. 277).)

The methodological quality to which I make reference relates to the care with which the Worcester group approached the critical issue of the patient's cooperativeness. A final quote from Shakow will make clear the precision that was accorded ratings for this control variable.

Another problem which troubles psychologists perennially is the part played by the cooperation or attitude of the subject. Almost all psychological tests and experiments require at least the passive participation of the subject. The data from such studies, except those directly investigating functioning at nonoptimal levels (a hazardous procedure, I must point out), carry the implication of having been collected under optimal conditions—external as well as internal. When there is suspicion that nonoptimal conditions are present, justifiable doubt about the validity of the findings arises. The argument may be offered that poverty in cooperativeness is intrinsic to schizophrenia; therefore, any attempt at the separa-

tion of its effects is at best academic. This thesis had validity to the extent that poor cooperation is intrinsic. The argument, however, runs into the difficulty of not making a distinction between the intrinsic effects of attitude and of other temporary or superficial interfering effects. In order to control for this factor in our studies, we consistently used an A to E rating scale which defined various levels of cooperation. The patients used in the studies reported fall mainly into the classes we labeled A and B cooperation, those showing either active interest in the task itself or active effort because of secondary interest (100, p. 276).

The effect of these twin factors of chronicity and cooperativeness undoubtedly served to produce markedly homogeneous groups in which the control of less relevant subject variables may have permitted those very aspects of subtype "style" to which Shakow has alluded to exercise some effect. Whether or not subtype composition would have an equivalent yield with an acute population is a valid question that can only be determined through empirical study. But I must confess that I am not sanguine about the success of such a venture.

My pessimism is induced by the results of a survey that was conducted under the aegis of the Duke Schizophrenia Project that Dr. Eliot H. Rodnick and I headed. I present these data showing the relationship between diagnostic subtypes and Phillips premorbid score ratings most tentatively because they are derived from a biased sample of cases—namely, those patients who participated in our research. It would have been far more impressive had we been able to gather comparable data on consecutive admissions of newly hospitalized schizophrenic patients; but unfortunately such data were not available to us.

Furthermore, I sound the cautionary note that the vagaries that inhere in psychiatric diagnosis all operate here. In exchange for this apologia, however, I would note that the hospital settings in which we worked were undoubtedly more typical of

the contemporary American psychiatric scene than Worcester State Hospital under the guiding hands of its outstanding professional group of the 1930's and 1940's. Thus the data are probably more representative of the typical field situation in which an investigator is likely to find himself. In figure 5, curves are presented showing the relationship of total Phillips premorbid scale scores to three classes of schizophrenic patients totaling 126 cases and 35 normal subjects. As can be seen, the degree of overlap of schizophrenic subtypes into the "good" premorbid category is so great that it insures behavioral heterogeneity. Such effects are less apparent with the chronic and acute undifferentiated schizophrenic subjects, but quite marked for the paranoid and miscellaneous groups. In effect, I am asserting that it is quite probable that a selection of paranoid, or catatonic, or hebephrenic subjects will distribute itself so widely along a "process-reactive" continuum that one should not expect to

secure consistent and statistically reliable differences over the wide range of laboratory tests that have been cited in this chapter. (Parenthetically, I would add that the bimodal distribution observed by Wittman is not as striking for the total Duke schizophrenic group despite selection procedures that were designed to insure good and poor premorbid subgroups.)

All told, the evidence suggests that greater reliance can be placed on the reduction of subject variability through selection based on those procedures that separate subjects on the basis of prognostic rather than symptom criteria.

Application of the Concept of a Process-Reactive Continuum to All Forms of Psychopathology

The final section of this paper is devoted to an issue of overriding conceptual significance. Is the process-reactive concept unique solely to schizophrenia or is it applicable to all forms of functional mental

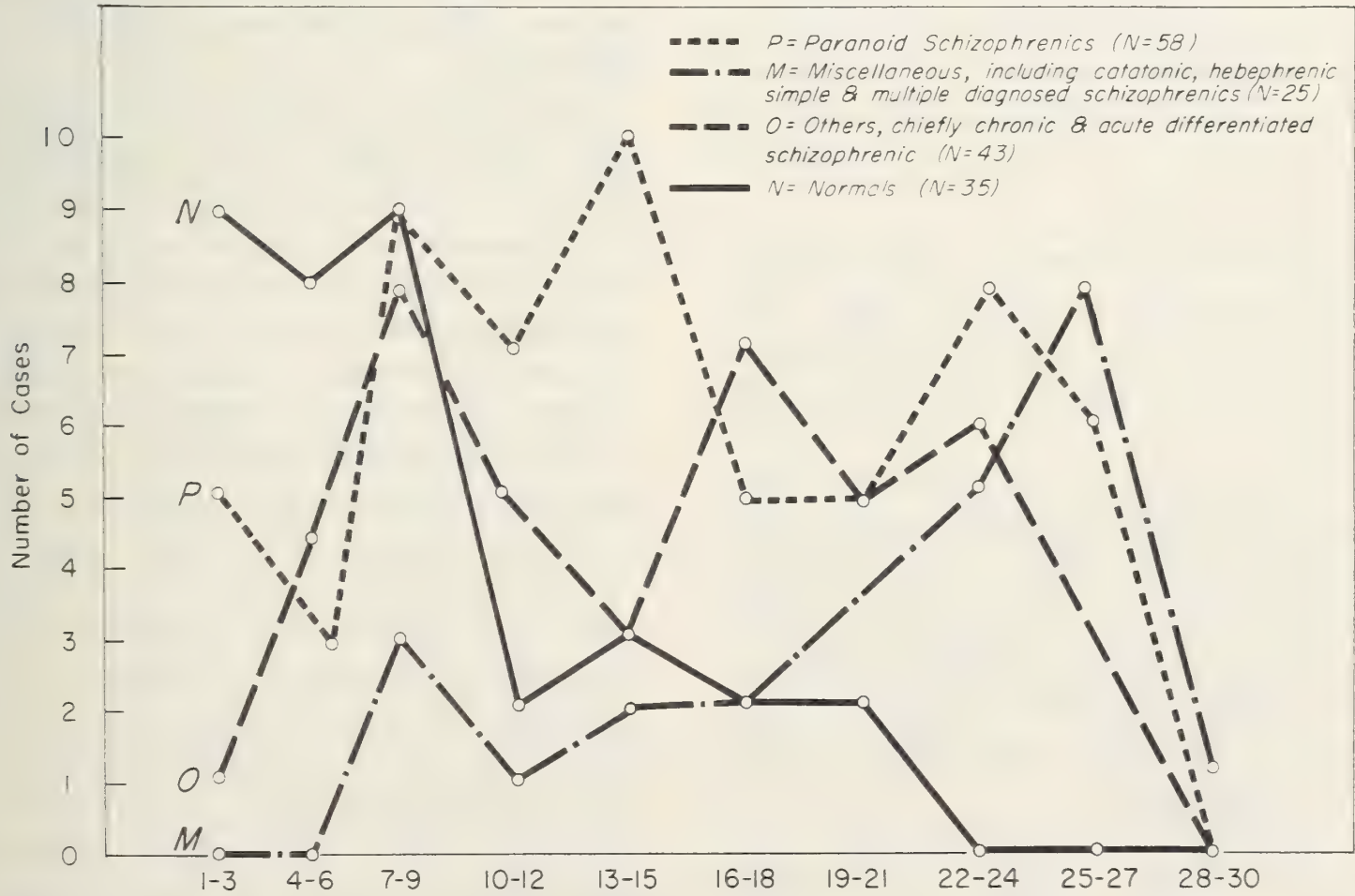


FIGURE 5.—Distribution of total premorbid scores (Phillips Scale) for schizophrenic diagnostic groups and normal subjects.
(From Duke Schizophrenia Project)

disorders? The minimal data available to us support the latter position, although we lack experimental studies that would provide a more determinate test of how extensively the concept can be generalized. The research of Zigler and Phillips on the relationship of social competence to symptom formation, diagnosis, and psychiatric outcome is most relevant to a consideration of this issue (111 through 117).

These investigators have traced the correlates of a social competence index that embraces six variables: age, intelligence, education, occupation, employment history, and marital status. Variables such as these bear, of course, a striking parallel to the earlier research on premorbid adequacy conducted by Phillips. Each variable is divided into three categories (scored 0-1-2 in increasing order of social competence), presumably representing points along a continuum of social adequacy. As described by Zigler and Phillips, the categories within each variable and their order from low to high social competence are as follows:

1. *Age*.—24 and below, 25-44, and 45 years and above.

2. *Intelligence*.—IQ's based on a standard intelligence test of 84 or less, 85-115, and 116 and above.

3. *Education*.—None or some grades including ungraded or special classes; completed grade school, some high school, or completed high school; and some college or more.

4. *Occupation*.—"The Dictionary of Occupational Titles" is used to assign a given occupation into one of the following groups: unskilled or semiskilled, skilled, and service; clerical and sales; professional and managerial. (The variable is left unscored for individuals not falling into one of the categories.)

5. *Employment history*.—Usually unemployed; seasonal, fluctuating, frequent shifts in employment or part-time employment; and regular employment.

6. *Marital status*.—Single; separated,

divorced, remarried, or widowed; and single continuous marriage.

Zigler and Phillips view mental disorders as representing a "continuous process in which the premorbid, initial, middle, and ultimate stages are meaningfully related." A consequence of this developmental position is their view that "the relationship of achieved level of maturity (defined in terms of premorbid social competence) to certain dimensions of psychopathology is not unique to schizophrenia but instead cuts across all forms of functional mental disorder" (117 p. 216). They find support for this position in their study of outcome in mental disorder in which schizophrenic and nonschizophrenic groups were compared (115). Using a total sample of 251 patients who had been admitted over a nine year period to Worcester State Hospital for a functional disorder, these investigators studied the relationship of social competence to hospital stay or hospital discharge, to length of institutionalization, to the frequency of rehospitalization and to the time elapsing between discharge and readmission.

Except for the failure to differentiate highs from lows on the last-named prognostic criterion, the high competence group systematically showed a higher frequency of discharge, a shorter period of institutionalization, and fewer instances of rehospitalization. A subsequent study, in which the form of treatment received by the patient was controlled, produced comparable findings, thus suggesting that the social competence variable may be a more significant predictor of recovery than is the treatment the patient receives.

Another type of evidence bearing on the generality of the process-reactive continuum requires a somewhat extended chain of inference. Levine (118) has observed that the relationship between Rorschach genetic scores of developmental maturity previously found to differentiate process and reactive schizophrenia (15, 92) can also serve to discriminate among non-

schizophrenic psychotics and is related to length of hospitalization. Thus a relationship, albeit a tenuous one, is suggested between levels of premorbid social competence, outcome in schizophrenic and nonschizophrenic psychosis, and level of maturity as expressed by Rorschach developmental scores.

Another empirical chain relating the process-reactive continuum to nonschizophrenic attributes is forged by research relating social attainment to reactions to stress in nonpathological subjects. Relevant findings include evidence that, even among normal men, degree of maladjustment is related to social adequacy (119); furthermore, normal subjects of lower levels of social attainment perform less adequately on psychomotor tasks run under failure-stress conditions (120); and higher social attainers show a greater adaptive response to a failure-induced stress including a less marked physiological reaction and a more realistic shift in goals following such failure (121).

Still another link is provided by the observation that social attainment is also related to the developmental level of perceptual functioning on the Rorschach Test (122).

And, finally, the study by Zigler and Phillips of the relationship of social competence to symptom expression in the functional disorders is particularly suggestive (117). Earlier studies had identified (114) a triadic cluster of symptoms that correlated with level of social competence. "Self-deprivation and turning against the self" (typically suicidal attempts and ideas, bodily complaints, tension headaches, self-depreciatory and depressive behaviors, etc.) were found to be positively related to a higher level of social competence, more favorable outcome, the presence of manic-depressive and psychoneurotic diagnoses and negatively related to the character disorders and schizophrenias. Symptoms of "self-indulgence and turning against

others" (e.g., maniacal outbursts, perversions, drinking, rape, assaultiveness, lying, etc.) were correlated with lower levels of social competence, predicted a less successful outcome and were positively related to character disorders and negatively related to manic-depressive, psychoneurotic, and schizophrenic diagnoses. Finally, a cluster of symptoms defined as "avoidance of others" (withdrawn, suspicious, hallucinatory, bizarre ideas, etc.) was also found in patients of low social competence, predicted a less successful course, and was positively related to schizophrenia while negatively related to the other three diagnostic groups.

When the relationships between symptom patterns and the process-reactive continuum in all four diagnostic groups were compared, the results suggested that such relationships transcended any single diagnostic group. Individuals with high symptom scores (defined as an emphasis on symptoms in the self-deprivation or turning against the self category) were found to have high social competence scores, whereas low social competence indicators in all four diagnostic groups were accompanied by low symptom scores. Zigler and Phillips concluded:

This finding, in conjunction with the earlier finding that both schizophrenic and nonschizophrenic patients with high social competence scores have a better prognosis than those with low scores, calls into question the heuristic value of the process-reactive distinction in schizophrenia. The implication here is that the process-reactive distinction is reducible to the social competence dimension, which is continuous in nature, and which is applicable not only to schizophrenia but to all of psychopathology. This approach makes of psychopathology a unitary phenomenon rather than a collection of discrete entities, each conceptualized in terms of unique dimensions and parameters * * * Although in general the conventional psychiatric categories are related to the level of social maturity attained one finds patients with both high and low social competence scores

in each of the conventional diagnostic categories. This less than perfect relationship may simply reflect what has long been suspected, namely, that in terms of the meaningful correlates of the diagnostic categories, the conventional diagnostic system is a relatively sterile one. It further suggests that the ability to predict such factors as outcome could be enhanced through the construction of a new diagnostic system based on a combination of premorbid social competence measures and related symptom dimensions (117, pp. 220-221).

I am not prepared, as are Zigler and Phillips, to dismiss our traditional psychiatric nomenclature. But the implications for research in these data are very clearly indicated. Most studies that have utilized a process-reactive separation for schizophrenics have not set forth the same distinction among the normal control or alternate psychiatric groups that have been used. I can point to the Duke Schizophrenia Project as a case in point. In the diagram below I have checked the groups used repeatedly in those studies. It is perfectly clear that Box 4 has not been a component part of these studies. And for good reason. Both the Phillips and the Elgin Scales have such a low ceiling that most nonschizophrenics (and virtually all normal subjects) tend to score at the more reactive end of the scale. Zigler and Phillips, in the study cited above, have noted this tendency for nonschizophrenic patients to have both higher social competence and higher symptom scores relative to the schizophrenic patients. But without Box 4, we cannot separate differences that are due to schizophrenic processes from those that are a function of social inadequacy.

Hopefully, the Social Competence Index may provide a more adequate method for separating patients of all diagnostic categories than has hitherto been available to us. A diligent search of patients may enable us to fill cell 4 in future studies with appropriate nonschizophrenic controls. But it will not be easy. The research of Zigler and Phillips is based upon case records garnered over a nine year period. At any given point in time, finding high and low social competence depressives, psychoneurotics, character disorders, and normals will be difficult. If we can manage to do so, then it will be profitable to initiate studies that go far beyond Zigler and Phillips' preoccupation with syndromes of symptoms. Studies to be initiated could include replications of those research investigations that have fruitfully differentiated between process-type and reactive-type schizophrenics.

A program of research into the biochemical, psychological, and social aspects of patient functioning could provide more definitive answers to a variety of questions. How adaptive under stressful and nonstressful conditions is the functioning of process-type, nonschizophrenic cases in relation to a comparable group of schizophrenic patients? Do such groups differ at varying levels of physiological functioning? To what extent do the basic patterns of socialization differ in such groups? Are both groups similarly responsive to different classes of social and nonsocial reinforcers? Do both groups show similar or different patterns of adaptation over time to new and stressful situations?

These are but a few of the many questions that could be asked if we were to provide, within our investigations, for groups of high and low social competence schizophrenic and nonschizophrenic patients with attendant normal controls. The potential productivity of such an approach is limitless. Were we to assume, for example, the continuity of premorbid, morbid, and postmorbid adaptation, then re-

DIAGNOSTIC CATEGORIES

		Schizophrenics	Normal controls
Social Adequacy	(Good (Reactive)	1 ✓	2 ✓
	(Poor (Process)	3 ✓	4

search with potentially and actually disordered children could be given comparable attention. Children who were already beginning to show failures in adaptation as well as others who came from environments that provide high base rates for cases of adult psychopathology could become subjects for comparable studies of cognition, language, learning, and perception to which I have alluded earlier. Beyond the heuristic value of such a program of research would be the important growth of attention to preventive aspects of disorder that these studies would inevitably entail. Such a venture warrants extended exploration.

Summary

This paper has reviewed some aspects of psychiatry's historical concern with the process-reactive distinction in schizophrenia, focusing upon a number of issues central to the concept: (1) Are process and reactive schizophrenia two separate and distinct disorders, or do both stem from a common disease entity? (2) Is the process-reactive distinction best viewed as a dichotomous typology or as a continuum? (3) Can process-schizophrenia be equated with an organic etiology and reactive schizophrenia with psychological origins?

Studies of rating scales designed to measure degrees of process-reactive status are reviewed. The relative efficacy of Kraepelinian nosology as opposed to the process-reactive distinction for reducing subject heterogeneity is considered. Finally, the extension of the concept to all forms of psychopathology is discussed with particular emphasis on measures of social competence.

REFERENCES

- (1) FRANK, J. "Psychoanalyse und Psychiatrie." *Samml. Psychoan. Aufsätze*, 1932, 99-102.
- (2) JASPERS, K. "Kausale and verständliche Zusammenhänge zwischen Schicksal und "Psychose bei der Dementia Praecox." *Z. Neurol. Psychiat.*, 1913, 14, 158.
- (3) BLEULER, E. "Textbook of Psychiatry." New York: Macmillan Co., 1924.
- (4) ZILBOORG, G., and HENRY, G. W. "A History of Medical Psychology." New York: copyright, 1941, by W. W. Norton & Co., Inc.
- (5) KRAEPELIN, E. "Clinical Psychiatry." New York: Macmillan Co., 1907.
- (6) KATZENELBOGEN, S. "Dementia Praecox." Formulation by Kraepelin, Bleuler, and Meyer. *Psychiat. Quart.*, 1942, 16, 439-453.
- (7) MEYER, A. "The Dynamic Interpretation of Dementia Praecox." *Amer. J. Psychol.*, 1910, 21, 385-403.
- (8) LANGFELDT, G. "The Prognosis in Schizophrenia." *Acta psychiat. Scand.*, 1956, suppl. 110.
- (9) ASTRUP, C., FOSSUM, A., and HOLMBOE, R. "Prognosis in Functional Psychoses." Springfield, Ill.: Charles C. Thomas, 1962.
- (10) SULLIVAN, HARRY S. "Schizophrenia as a Human Process." New York: W. W. Norton & Co., copyright (©), 1962, by the William Alanson White Psychiatric Foundation.
- (11) WIENER, H. "Diagnosis and Symptomatology." In *Schizophrenia: A Review of the Syndrome*. Bellak, L. (editor). New York: Logos Press, 1958, 107-173.
- (12) KANTOR, R. E., WALLNER, J. M., and WINDER, C. L. "Process and Reactive Schizophrenia." *J. consult. Psychol.*, 1953, 17, 157-162.
- (13) HUSTON, P. E., and PEPERNIK, M. C. "Prognosis in Schizophrenia." In *Schizophrenia: A Review of the Syndrome*. Bellak, L. (editor). New York: Logos Press, 1958, pp. 531-554.
- (14) WITTMAN, P. "A Scale for Measuring Prognosis in Schizophrenic Patients." *Elgin State Hospital Papers*, 1941, 4, 20-33.
- (15) BECKER, W. C. "The Relation of Severity of Thinking Disorder to the Process-Reactive Concept of Schizophrenia." Unpublished doctoral dissertation, Stanford University, 1955.
- (16) PHILLIPS, L. "Case History Data and Prognosis in Schizophrenia." *J. nerv. ment. Dis.*, 1953, 117, 515-525.
- (17) McDONOUGH, M. M. "Critical Flicker Frequency and the Spiral After-Effect With Process and Reactive Schizophrenics." *J. consult. Psychol.*, 1960, 24, 150-155.
- (18) CHAPMAN, L. J., DAY, D., and BURSTEIN, A. "The Process-Reactive Distinction and Prognosis in Schizophrenia." *J. Nerv. ment. Dis.*, 1961, 133, 383-391.
- (19) FINE, H. J., and ZIMET, C. N. "Process-Reactive Schizophrenia and Genetic Levels of Perception." *J. abnorm. soc. Psychol.*, 1959, 59, 83-86.
- (20) KING, G. "Differential Autonomic Responsiveness in the Process-Reactive Classification of Schizophrenia." *J. abnorm. soc. Psychol.*, 1958, 56, 160-164.

- (21) ZUCKERMAN, M., and GROSZ, H. J. "Contradictory Results Using the Mecholy Test to Differentiate Process and Reactive Schizophrenia." *J. abnorm. soc. Psychol.*, 1959, 59, 145-146.
- (22) LORR, M., WITTMAN, P., and SCHANBERGER, W. "An Analysis of the Elgin Prognostic Scale." *J. clin. Psychol.*, 1951, 7, 260-263.
- (23) BECKER, W. C. "The Process-Reactive Distinction: A Key to the Problem of Schizophrenia." *J. nerv. ment. Dis.*, 1959, 129, 442-449.
- (24) BECKER, W. C. "A Genetic Approach to the Interpretation and Evaluation of the Process-Reactive Distinction in Schizophrenia." *J. abnorm. soc. Psychol.*, 1956, 53, 229-236.
- (25) STEFFY, R. A., and BECKER, W. C. "Measurement of the Severity of Disorder in Schizophrenia by Means of the Holtzman Inkblot Test." *J. consult. Psychol.*, 1961, 25, 555.
- (26) STEFFY, R. A., and BECKER, W. C. "Extended Report: Measurement of the Severity of Disorder in Schizophrenia by Means of the Holtzman Inkblot Test." Unpublished manuscript. University of Illinois, Urbana, 1961.
- (27) WITTMAN, PHYLLIS, and STEINBERG, L. "Followup of an Objective Evaluation of Prognosis in Dementia Praecox and Manic-Depressive Psychoses." *Elgin State Hospital Papers*, 1944, 5, 216-227.
- (28) WITTMAN, PHYLLIS. "Followup on Elgin Prognostic Scale Results." *Illinois Psychiat. J.*, 1944, 4, 56-59.
- (29) SEIDEL, CLAUDEINE. "The Relationship Between Klopfer's Rorschach Prognostic Rating Scale and Phillips' Case History Prognostic Rating Scale." *J. consult. Psychol.*, 1960, 24, 46-49.
- (30) RODNICK, E. H., and GARMEZY, N. "An Experimental Approach to the Study of Motivation in Schizophrenia." In Jones, M. R. (editor), *Nebraska Symposium on Motivation*. Lincoln: University of Nebraska Press, 1957, pp. 109-184.
- (31) GARMEZY, N., FARINA, A., and RODNICK, E. H. "Reliability Studies of the Phillips Scale." Unpublished data. Duke Schizophrenia Research Project. Durham, N. C., 1960.
- (32) ULLMANN, L. P., and GIOVANNONI, JEANNE M. "The Development of a Self-Report Measure of the Process-Reactive Continuum." *J. nerv. ment. Dis.*, 1964, 138, 38-42.
- (33) JENKINS, R. L., and GUREL, L. "Predictive Factors in Early Release." *Mental Hospitals*, 1959, 11-14.
- (34) FARINA, A., and WEBB, W. W. "Premorbid Adjustment and Subsequent Discharge." *J. nerv. ment. Dis.*, 1956, 124, 612-613.
- (35) FARINA, A., GARMEZY, N., ZALUSKY, M., and BECKER, J. "Premorbid Behavior and Prognosis in Female Schizophrenic Patients." *J. consult. Psychol.*, 1962, 26, 56-60.
- (36) FARINA, A., GARMEZY, N., and BARRY, HERBERT, III. "Relationship of Marital Status to Incidence and Prognosis of Schizophrenia." *J. abnorm. soc. Psychol.*, 1963, 67, 624-630.
- (37) SOLOMON, L., and ZLOTOWSKI, M. "The Relationship Between the Elgin and the Phillips Measures of Process-Reactive Schizophrenia." *J. nerv. ment. Dis.*, 1964, 138, 32-37.
- (38) HIGGINS, J. "The Concept of Process-Reactive Schizophrenia." *Criteria and Related Research*. *J. nerv. ment. Dis.*, 1964, 138, 9-25.
- (39) HERRON, W. G. "The Process-Reactive Classification of Schizophrenia." *Psychol. Bull.*, 1963, 59, 329-343.
- (40) GARMEZY, N., and RODNICK, E. H. "Premorbid Adjustment and Performance in Schizophrenia." *J. nerv. ment. Dis.*, 1959, 129, 450-466.
- (41) SCHAEFER, S. M. J. "MMPI Profiles of Good and Poor Premorbid Schizophrenic Patients." Summa cum laude thesis. University of Minnesota, Minneapolis, 1965.
- (42) WELSH, G. S. "Factor Dimensions A and R." In Welsh, G. S., and Dahlstrom, W. G. (editors). *Basic Readings on the MMPI in Psychology and Medicine*. Minneapolis: University of Minnesota Press, 1946.
- (43) BARRON, F. "An Ego-Strength Scale Which Predicts Response to Psychotherapy." *J. consult. Psychol.*, 1953, 17, 327-333.
- (44) SCHOFIELD, W. "Changes in Responses to the MMPI Following Certain Therapies." *Psychol. Monogr.*, 1950, 64, No. 5. (Whole No. 311.)
- (45) SCHOFIELD, W. "A Further Study of the Effects of Therapies on MMPI Responses." *J. abnorm. soc. Psychol.*, 1953, 48, 67-77.
- (46) JENKINS, W. L. "The MMPI Applied to the Problems of Prognosis in Schizophrenia." Unpublished Ph. D. thesis. University of Minnesota, Minneapolis, 1952.
- (47) CHURCH, J. C. "Relationship Between Premorbid History and MMPI Profiles of Schizophrenics." Master's thesis. University of North Carolina, Chapel Hill, 1963.
- (48) WITTMAN, P. "A Scale for Measuring Prognosis in Schizophrenic Patients." *Elgin State Hospital Papers*, 1941, 4, 20-33.
- (49) GARMEZY, N., CLARKE, A. R., and STOCKNER, C. "Child-rearing Attitudes of Mothers and Fathers as Reported by Schizophrenic and Normal Patients." *J. abnorm. soc. Psychol.*, 1961, 63, 176-182.
- (50) HARRIS, J. G., Jr. "Size Estimation of Pictures as a Function of Thematic Content for Schizophrenic and Normal Subjects." *J. Pers.*, 1957, 25, 651-672.
- (51) BLEKE, R. C. "Reward and Punishment as Determiners of Reminiscence Effects in Schizophrenic and Normal Subjects." *J. Pers.*, 1955, 23, 479-498.

- (52) HERRON, W. G. "Abstract Ability in the Process-Reactive Classification of Schizophrenia." *J. gen. Psychol.*, 1962, 67, 147-154.
- (53) MORIARTY, D., and KATES, S. L. "Concept Attainment of Schizophrenics on Materials Involving Social Approval and Disapproval." *J. abnorm. soc. Psychol.*, 1962, 65, 355-364.
- (54) ZUCKERMAN, M., and GROSZ, H. J. "Contradictory Results Using the Mecholyl Test to Differentiate Process and Reactive Schizophrenia." *J. abnorm. soc. Psychol.*, 1959, 59, 145-146.
- (55) JUDSON, A. J., and KATAHN, M. "The Relationship of Autonomic Responsiveness to Process-Reactive Schizophrenia and Abstract Thinking." *Psychiat. Quart.*, 1963, 37, 19-24.
- (56) FARINA, A. "Patterns of Role Dominance and Conflict is the Interaction of Parents of Schizophrenic Patients." Ph. D. thesis. Duke University, Durham, N. C., 1958.
- (57) SLOTNICK, LEE K. "The Effects of the Interpersonal Interaction Situation Upon Abstract Thinking Performance in Schizophrenia." Unpublished Ph. D. thesis. University of Illinois, Urbana, Ill., 1963.
- (58) TUTKO, T. A., and SPENCE, JANET T. "The Performance of Process and Reactive Schizophrenics and Brain Injured Subjects on a Conceptual Task." *J. abnorm. soc. Psychol.*, 1962, 65, 387-394.
- (59) CAMERON, N. "Reasoning, Regression and Communication in Schizophrenics." *Psychol. Monogr.*, 1938, 58, 1. (Whole No. 221.)
- (60) EPSTEIN, S. "Overinclusive Thinking in a Schizophrenic and a Control Group." *J. consult. Psychol.*, 1953, 17, 384-388.
- (61) MCGAUGHRAN, L. S., and MORAN, L. J. "Conceptual Level" versus "Conceptual Area" Analysis of Object Sorting Behavior of Schizophrenic and Nonpsychiatric Groups." *J. abnorm. soc. Psychol.*, 1956, 52, 43-50.
- (62) PAYNE, R. W., and HEWLETT, J. H. G. "Thought Disorder in Psychiatric Patients." In H. J. Eysenck (editor). *Experiments in Personality*. Vol. 11. London: Routledge & Kegan Paul, 1960.
- (63) JASPERS, K. "General Psychopathology." Chicago: University of Chicago Press, 1963 English translation of "Allgemeine Psychopathologie," 7th edition. From the German by J. Hoenig and Marian W. Hamilton.
- (64) TYRELL, D. J., STRUVE, F. A., and SCHWARTZ, M. L. "A Methodological Consideration in the Performance of Process and Reactive Schizophrenics on a Test of Organic Brain Pathology." *J. clin. Psychol.*, 1965, 21, 254-256.
- (65) KETY, S. S. "Biochemical Theories of Schizophrenia." *Science*, 1959, 129, 1528-1532, 1590-1596, 3362-3363.
- (66) HORWITT, M. K. "Fact and Artifact in the Biology of Schizophrenia." *Science*, 1956, 124, 429-430.
- (67) ROSENTHAL, D. "Some Factors Associated With Concordance and Discordance With Respect to Schizophrenia in Monozygotic Twins." *J. nerv. ment. Dis.*, 1959, 129, 1-10.
- (68) SLATER, E. "Psychotic and Neurotic Illnesses in Twins." London: Her Majesty's Stationery Office, 1953.
- (69) ROSENTHAL, D. "The Genain Quadruplets." New York: Basic Books, 1963.
- (70) WYNNE, L. C., DAY, JULIANA, HIRSCH, S. T., and RYCKOFF, I. "The Family Relations of a Set of Monozygotic Quadruplet Schizophrenics." Congress Report of the Second International Congress for Psychiatry. Zurich, Switzerland, 1957, vol. II, pp. 43-49.
- (71) MEEHL, P. E. "Schizotaxia, Schizotypy, Schizophrenia." *Amer. Psychol.*, 1962, 17, 827-838.
- (72) GERARD, RALPH W. "The Organic Position on Etiology." In *Interpreting Approaches to Mental Disease*. Kruse, H. D. (editor). New York: Hoeber-Harper, 1957.
- (73) DUNN, W. L., Jr. "Visual Discrimination of Schizophrenic Subjects as a Function of Stimulus Meaning." *J. Pers.*, 1954, 23, 48-64.
- (74) DUNHAM, R. M. "Sensitivity of Schizophrenics to Parental Censure." Unpublished Ph. D. thesis. Duke University, Durham, N.C., 1959.
- (75) KREINIK, PHYLLIS S. "Parent-Child Themes and Concept Attainment in Schizophrenia." Unpublished Ph. D. thesis. Duke University, Durham, N.C., 1959.
- (76) BAXTER, J. C., and BECKER, J. "Anxiety and Avoidance Behavior in Schizophrenics in Response to Parental Figures." *J. abnorm. soc. Psychol.*, 1962, 64, 432-437.
- (77) BAXTER, J. C., BECKER, J., and HOOKS, W. "Defensive Style in the Families of Schizophrenics and Controls." *J. abnorm. soc. Psychol.*, 1963, 66, 512-518.
- (78) GOODMAN, D. "Performance of Good and Poor Premorbid Male Schizophrenics as a Function of Paternal Versus Maternal Censure." *J. abnorm. soc. Psychol.*, 1964, 69, 550-555.
- (79) GARMEZY, N., FARINA, A., and RODNICK, E. H. "The Structured Situational Test: A Method for Studying Family Interaction in Schizophrenia." *Amer. J. Orthopsychiat.*, 1960, 30, 445-452.
- (80) LEVY, D. M. "Maternal Overprotection." New York: Columbia University Press, 1943.
- (81) LEWIS, N. D. C. "Research in Dementia Praecox." The National Committee for Mental Hygiene, 1936.
- (82) SILVERMAN, J. "Psychological Deficit Reduction in Schizophrenia Through Response-Contingent Noxious Reinforcement." *Psychol. Rep. Monogr. Suppl.*, 1963, 2-V13, 187-210.

- (83) DEWOLFE, A. S. "The Effect of Affective Tone on the Verbal Behavior of Process and Reactive Schizophrenics." *J. abnorm. soc. Psychol.* 1962, 64, 450-455.
- (84) DOECKI, P. R., POLIDORO, L. G., and CROMWELL, R. L. "Commonality and Stability of Word Association Responses in Good and Poor Premorbid Schizophrenics." *J. abnorm. Psychol.*, 1965, 70, 312-316.
- (85) DRAGUNS, J. G. "Responses to Cognitive and Perceptual Ambiguity in Chronic and Acute Schizophrenics." *J. abnorm. soc. Psychol.*, 1963, 66, 24-30.
- (86) REISMAN, J. M. "Motivational Differences Between Process and Reactive Schizophrenics." *J. Pers.*, 1960, 28, 12-25.
- (87) ALVAREZ, R. R. "A Comparison of the Preferences of Schizophrenic and Normal Subjects for Rewarded and Punished Stimuli." Unpublished Ph. D. thesis. Duke University, Durham, N.C., 1957.
- (88) DEVAULT, S. H. "Physiological Responsiveness in Reactive and Process Schizophrenia." Unpublished Ph. D. thesis. East Lansing, Michigan State University, 1955.
- (89) GARMEZY, N. "The Prediction of Performance in Schizophrenia." In *Schizophrenia*. Hoch, P., and Zubin, J. (editors). New York: Grune & Stratton, 1966.
- (90) FARINA, A. "Patterns of Role Dominance and Conflict in Parents of Schizophrenic Patients." *J. abnorm. soc. Psychol.*, 1960, 61, 31-38.
- (91) SOLOMON, L., KLEIN, E., and HERFORD, T. "Process-Reactive Schizophrenia, Personality Factors and Interpersonal Submission." Paper presented at Eastern Psychological Association meetings, New York City, 1963.
- (92) ZIMET, C. N., and FINE, H. J. "Perceptual Differentiation, and Two Dimensions of Schizophrenia." *J. nerv. ment. Dis.*, 1959, 129, 435-441.
- (92a) SILVERMAN, J. "The Problem of Attention In Research and Theory In Schizophrenia." *Psychol Rev.*, 1964, 71, 352-379.
- (93) SILVERMAN, J. "The Scanning-Control Mechanism and 'Cognitive Filtering' in Paranoid and Nonparanoid Schizophrenia." *J. consult. Psychol.*, 1964, 28, 385-393.
- (94) MAYER-GROSS, W., SLATER, E., and ROTH, M., "Clinical Psychiatry." Baltimore: Williams & Wilkins Co., 1960.
- (95) HENDERSON, SIR DAVID, and BATCHELOR, IVOR R. C. "Henderson and Gillespie's Textbook of Psychiatry." London: Oxford University Press, 1962.
- (96) HOSKINS, R. G. "The Biology of Schizophrenia." New York: W. W. Norton Co., 1946.
- (97) HUSTON, P. E., and SHAKOW, D. "Studies of Motor Function in Schizophrenia: III. Steadiness." *J. gen. Psychol.*, 1946, 34, 119-126.
- (98) HUSTON, P. E., SHAKOW, D., and RIGGS, L. A. "Studies of Motor Function in Schizophrenia." *J. gen. Psychol.*, 1937, 16, 39-82.
- (99) HUSTON, P. E., and SHAKOW, D. "Learning in Schizophrenia. I. Pursuit learning." *J. Pers.*, 1948, 17, 52-74.
- (100) SHAKOW, D. "Psychological Deficit in Schizophrenia." *Behav. Sci.*, 1963, 8, 4, 275-305.
- (101) HAVENER, P. H., and IZARD, C. E. "Unrealistic Self-Enhancement in Paranoid Schizophrenics." *J. consult. Psychol.*, 1962, 26, 65-68.
- (102) MANDL, BILLIE, S. T. "An Investigation of Rigidity in Paranoid Schizophrenics as Manifested in a Perceptual Task." Unpublished Ph. D. thesis. Purdue University, 1954.
- (103) NEWBERRY, J. J. "Perceptual Thresholds of Paranoid Schizophrenics as a Function of Affectively Toned Words Having Interpersonal Implications." Unpublished Ph. D. thesis. Northwestern University, 1954.
- (104) SHAKOW, D. "Segmental Set." *Arch. gen. Psychiat.*, 1962, 6, 1-17.
- (105) SILVERMAN, J. "The Perceptual Control of Stimulus Intensity in Paranoid and Nonparanoid Schizophrenia." *J. nerv. ment. Dis.*, 1964, 139, 545-549.
- (106) HARTMAN, A. M. "The Apparent Size of After-Images in Delusional and Nondelusional Schizophrenics." *Amer. J. Psychol.*, 1962, 75, 587-595.
- (107) RAUSCH, H. L. "Perceptual Constancy in Schizophrenia." *J. Pers.*, 1952, 21, 176-187.
- (108) SEXTON, M. C. "The Autokinetic Test: Its Value in Psychiatric Diagnosis and Prognosis." *Amer. J. Psychiat.*, 1945, 102, 399-402.
- (109) VOTH, A. C. "An Experimental Study of Mental Patients Through the Autokinetic Phenomenon." *Amer. J. Psychiat.*, 1947, 103, 793-805.
- (110) VOTH, H. M. "Choice of Illness." *Arch. gen. Psychiat.*, 1962, 6, 149-156.
- (111) ZIGLER, E., and PHILLIPS, L. "Social Effectiveness and Symptomatic Behaviors." *J. abnorm. soc. Psychol.*, 1960, 61, 231-238.
- (112) ZIGLER, E., and PHILLIPS, L. "Case History Data and Psychiatric Diagnosis." *J. consult. Psychol.*, 1961, 25, 458.
- (113) ZIGLER, E., and PHILLIPS, L. "Psychiatric Diagnosis: A Critique." *J. abnorm. soc. Psychol.*, 1961, 63, 69-75.
- (114) ZIGLER, E., and PHILLIPS, L. "Psychiatric Diagnosis and Symptomatology." *J. abnorm. soc. Psychol.*, 1961, 63, 69-75.
- (115) ZIGLER, E., and PHILLIPS, L. "Social Competence and Outcome in Psychiatric Disorder." *J. abnorm. soc. Psychol.*, 1961, 63, 264-271.

(116) PHILLIPS, L., and ZIGLER, E. "Social Competence: The Action-Thought Parameter and Vicariousness in Normal and Pathological Behaviors." *J. abnorm. soc. Psychol.*, 1961, 63, 137-146.

(117) ZIGLER, E., and PHILLIPS, L. "Social Competence and the Process-Reactive Distinction in Psychopathology." *J. abnorm. soc. Psychol.*, 1962, 65, 215-222.

(118) LEVINE, D. "Rorschach Genetic Level and Mental Disorder." *J. proj. Tech.*, 1959, 23, 436-439.

(119) FINE, H. J., FULKERSON, S. C., PHILLIPS, L. "Maladjustment and Social Attainment." *J. abnorm. soc. Psychol.*, 1955, 50, 33-35.

(120) FEFFER, H., and PHILLIPS, L. "Social Attainment and Performance Under Stress." *J. Pers.*, 1953, 22, 284-297.

(121) GERARD, D. L., and PHILLIPS, L. "Relation of Social Attainment to Psychological and Adrenocortical Reactions to Stress." *Arch. neurol. Psychiat.*, 1953, 69, 352-354.

(122) LANE, J. E. "Social Effectiveness and Developmental Level." *J. Pers.*, 1955, 23, 274-284.

APPENDIX A

ELGIN PROGNOSTIC SCALE

(Becker revision—abbreviated scale items are denoted by asterisks)

The definitions for the subscales of the Elgin Prognostic Scale as modified by Becker are given below. Items A through O are rated on the basis of the anamnesis data. Items P through T are rated on the basis of the presenting clinical symptoms.

*A. Defects of Interest Versus Definite Display of Interest

0. Keen ambitious interest in some of the following: Home, family, friends, work, sports, arts, pets, gardening, social activities, music, dramatics.
2. Moderate degree of interest in several activities including social gatherings, sports, music, opposite sex, etc.
4. Mild interest in a few things such as job, family, quiet social gatherings. The interest is barely sustaining.
6. Withdrawn and indifferent toward life interests of average individual. No deep interests of any sort.

*B. Insidious Versus Acute Onset

0. Development over a period of 0 to 1 month with sudden, dramatic divorcement from more or less commonplace living.

1. Development over a period of 2 to 4 months with marked personality changes from relatively commonplace living.
2. Development over a period of 5 to 7 months with moderate personality changes. May be some accenting of previous trends, but changes also.
3. Changes have taken place over a period of 8 to 12 months, with noticeable personality modifications, but primarily an accenting of existing trends.
4. Slow development of symptoms, but possible to detect personality changes in 2 years prior to onset.
6. Very slow development of symptoms so that final disorder appears as an exaggeration of already strongly accentuated personality traits. Indications even prior to adolescence.

*C. Shut-in Personality

General.—The psychotic condition is simply an exaggeration of the peculiar type of personality shown all through childhood. Stormy childhood often with overprotection and anxiety, a difficult adolescence characterized by inability to get along with and mix with other children. Constitutional apparently rather than product of specific environment.

0. Apparently normal childhood, little evidence of shyness, unusual difficulty, or else unusual behavior is attributable to environmental factors.
1. Only mildly this way, but some resemblance to pattern.
3. Moderately the picture described above.
5. Very much as described above.

D. Schizothymic Versus Syntononic Personality

0. Very sociable, fond of people and social gatherings; many friends, active in groups and sports, participates in life of his community.
2. Moderately sociable, likes people and social gatherings, but doesn't go far out of way to meet people.
3. Mildly shy, mildly sociable. Will interact when the situation presents itself. Prefers association in family group as a rule.
4. Moderately shy, retiring, etc. More concerned with ideas than people.
6. Very seclusive, shy, retiring, mixes little with others. Few if any close friends. Interested in ideas rather than people. Passive, an onlooker at life rather than an active participant. Poor "bite on life."

E. Range of Interests

0. Wide and varied interests, keen bite on life and its opportunities, forward and interested in making adaptations to daily life in many spheres.

2. Moderate breadth of interests, interested in making adaptations to daily life, but does not go out of way to seek new opportunities.
4. Moderate restriction of interests. Narrow goals, but some detectable variety of interests within a narrow orientation.
6. Inadequate interest in varied problems of life, rigid, narrow goals or interests, circumscribed activities because of the narrow range of interests.

F. Constitutional Bias

0. A healthy, strong, energetic, physical and mental makeup that makes the interplay between heredity and environmental influence during childhood a satisfactory one.
2. Suggestions of defects in physical and mental stamina occasionally observed. Not at all marked. Perhaps proneness to repeated illness in childhood.
4. Regarded from early childhood as different, queer or odd; perhaps associated with some real defect or handicap—physical, such as deformity, or speech defect, but more often only an imaginary defect of personality.

G. Low Energy Tone

0. Very strong drive, keen active, and alert interest and ambition shown in school, social and work spheres. Good grasp on life, liked life and had energy enough to enjoy it. Outgoing and adequate in meeting life.
2. Moderately adequate drive, interest, energy as described above.
4. Moderately inadequate energy tone. Tends toward submissive, passive reactions. Shows some potential to face life's problems, but would rather avoid them than expend the necessary energy.
6. Submissive, inadequate, passive reactions, weak grasp on life, does not go out to meet life's problems, does not participate actively but passively accepts his lot without having the energy to help himself.

H. Asthenic Build

0. Large, barrel-shaped trunk, with relatively short legs, and arms; shield-shaped face, short, broad head upon a thick neck, set well down between shoulders.
2. Athletic build. Balanced weight, good musculature, head shape, etc., intermediate to 0 and 4.
4. Long, slender extremities with relatively small, narrow trunk, egg-shaped face; elongated narrow head on a tall, slender neck.

**I. Heterosexual Contact*

0. Purposefully contacts the other sex, dates frequently, makes successful effort to be attractive in manner, dress, accessories, etc., so as to be popular with girls (boys).
2. Dates when situation affords. Maybe marries, but to have difficulties in compatibility. Wants to interact with other sex, has some techniques, but not completely successful.
3. If married, apt to divorce or separate. Generally this is rated as a midpoint between 2 and 4.
4. Moderate lack of heterosexual contact. Tends to avoid dates, dances, but has on occasion participated in same. Might think he (she) would like to marry someday, but has little enthusiasm for it.
6. No association with the opposite sex. Never had any dates. Avoids dances and social gatherings which require the intermingling of boys and girls.

J. Marked Academic Interests Versus Active Interests in Sports

0. An active interest in sports, participates in baseball, basketball, tennis, football or other sports. A solitary sport such as swimming or golf is not so important unless the patient plays or swims with others rather than himself.
1. Moderate interests in sports and other interests.
2. Mild interest in sports, mild interest in study.
3. Moderate interest in study—without other interests.
4. Fond of study, works diligently at school and excels in this field, associated with inadequacy in sports and social fields, a grind without the ambition or drive in work and play to equal his achievements as a student.

**K. Careless Indifference Versus Worrying, Self-Conscious Type*

0. Subjectively sensitive, critical of self, preoccupied with own conflicts, but shows little of the extreme, bizarre, unusual, mysterious or socially unacceptable in behavior.
2. Some concern and preoccupation with difficulties—a moderate position to 0 and 4.
4. Withdrawal and disinterest in social surroundings, careless of social requirements, given to daydreaming and eccentricity, dirty, dishevelled appearance, profane language, unacceptable habits.

**L. Exclusive Stubborn Traits Versus Insecurity and Inferiority Feelings*

0. Timid, lacks self-confidence, feels insecure and inferior. Very sensitive and critical of self; feels certain problems in life but participates and

does not accept his lot passively or without regret and struggle.

- 1. Moderately like 0.
- 2. Neither timid nor stubborn.
- 3. Moderately stubborn.
- 4. Complete withdrawal from surroundings and interests, inadequate in meeting life but stubborn and opinionated, refuses to change, even if suggested, to achieve a more adequate adjustment. Opinionated and egocentric.

M. Toxicity of Exhaustion

- 0. History of illness, disease, or exhaustion closely associated with the onset of psychotic symptoms.
- 1. Illness present, not severe, but related to onset. Less severe exhaustion.
- 2. Poor health—but not requiring bed.
- 3. Fair health—a little run down.
- 4. Excellent health history, health in no sense an etiologic factor in the development of psychosis.

*N. Precipitating Conditions (Situational Reaction)

- 0. A strong relationship between onset of symptoms and situational problems that would require definite and continued effort to adjust satisfactorily; i.e., death, failure, loss, interpersonal strife. The average person would definitely try to flee such a situation rather than attempt to change it.
- 1. Marked stresses related to onset, but not as severe as 0.
- 2. Moderate stresses related to onset such as financial problems, interpersonal discord, etc., which would cause considerable worry to the average individual.
- 3. Mild stresses that the average person would react to in some way but which would not usually lead to a breakdown.
- 4. Onset of psychotic symptoms not related to any disturbance or difficulty in the patient's situation—or a disturbance of such a trivial nature that it would be ignored or quickly forgotten by the average person.

*O. Duration of Psychosis

- | | |
|-------------------|--------------------|
| 0. Under 2 months | 4. 10 to 12 months |
| 1. 2 to 4 months | 5. 1 to 2 years |
| 2. 4 to 6 months | 6. 2 to 3 years |
| 3. 6 to 8 months | 7. Over 3 years |

The following scales are rated from the present-ing clinical picture:

P. Inadequate Affect Versus Emotional Instability or Appropriate Affect

- 0. Adequate or overly demonstrative affective ex-pression. This includes appropriate expression and manic-depressive aspects in which there is a facile display of emotion.

- 2. Moderately inadequate affect. Tends to be rigid, dull, or slightly inappropriate. Only moderate responsiveness to emotional stimulation.
- 4. Markedly inadequate, inappropriate, rigid, or dull affect. Emotional life expressed is at odds with behavior or strikingly inappropriate. Little reaction to stimulation of any strength.

*Q. Hebephrenic Symptoms

Extreme indifference, complete divorce between ideas and affect; extreme carelessness in appearance and reaction with untidiness; in some cases, silly behavior, often silly laughter with appropriate stimulation.

- | | |
|------------------------|---------------------------|
| 0. Not as above | 3. Markedly as above |
| 1. Mildly as above | 4. Very markedly as above |
| 2. Moderately as above | |

R. Ideas of Influence

Patient feels that someone or something is direct-ing his actions, thoughts, or speech. Some outside influence forces him to do things even against his own will.

(Rate 0-4 as in scale Q)

*S. Physical Interpretation of Delusions

The patient has certain feelings (possibly halluci-nations) that are linked up with definite delusional ideas; for instance, that there is a snake in his stomach, that food passes right through his body, that someone is passing electrical currents through his body; that the food he eats is poisoned, etc.

(Rate 0-4 as in scale Q)

T. Atypical Symptoms

Manic or depressive features mixed with the schizophrenic picture. Display of appropriate affect, overtalkative, distractive, facetious, display of in-terest in other patients, desire to help humanity in general, depression, feelings of sin or guilt, psy-chomotor retardation, anxiety, crying.

- 0. Very markedly atypical picture, shows many of the above features with considerable strength of affect.
- 1. Markedly atypical picture.
- 2. Moderately atypical picture, less intensity of features shown.
- 3. Mildly atypical picture, unusual features minimal or lacking in intensity.
- 4. Lacking atypical features.

If data utilized in filling out scale has been derived mainly from interview with patient, rate (0-4) the following:

- | | |
|--------------------------------------|-------|
| 1. Degree of patient cooperativeness | _____ |
| 2. Reliability of information | _____ |
| 3. Process reaction | _____ |

APPENDIX B

THE PHILLIPS SCALE OF PREMORBID ADJUSTMENT IN SCHIZOPHRENIA

(Modified with descriptive criteria by Farina and Garmezy for use with male and female patients)

I. PREMORBID HISTORY

A. Recent Sexual Adjustment

(NOTE.—Score as sexual contact; when information is not explicitly given, use inference to get at this actual sexual behavior.)

1. Stable heterosexual relation and marriage 0
2. Continued heterosexual relation and marriage but unable to establish home 1
3. Continued heterosexual relation and marriage broken by permanent separation 2
4. (a) Continued heterosexual relation and marriage but with low sexual drive 3
(NOTE.—If only informant is mother, don't score sexual adjustment. Prorate from rest of Premorbid History section. Look here for evidences of frigidity, distaste, avoidance, infrequency. Don't score on matters of technique.)
- (b) Continued heterosexual relation with deep emotional meaning but emotionally unable to develop it into marriage 3
(NOTE.—This must involve actual physical contact. Petting behavior is acceptable here. *Mutuality* of feeling is not necessary, but sexual behavior is, i.e., no adoration from afar.)
5. (a) Casual but continued heterosexual relations, i.e., "affairs" but nothing more 4
(NOTE.—"Casual" here implies lack of emotional meaning, although sexual behavior is consistent and regular.)
- (b) Homosexual contacts with lack of or chronic failure in heterosexual experiences 4
6. (a) Occasional casual heterosexual or homosexual experiences with no deep emotional bond 5
(NOTE.—This differs from 5(a) on the dimension of frequency. Contacts less often here.)
- (b) Solitary masturbation with no active attempt at homosexual or heterosexual experiences 5
7. No sexual interest in either men or women .. 6

B. Social Aspects of Sexual Life During Adolescence and Immediately Beyond

1. Always showed a healthy interest in the opposite sex—with a "steady" during adolescence 0

(NOTE.—"Steady" implies the exclusiveness of the dating relationship [neither partner dates anyone else] as well as frequency and emotional attachment.)

2. Started dating regularly in adolescence 1
(NOTE.—This implies twosomeness, pairing off into couples, as distinguished from 3, below)
3. Always mixed closely with boys and girls 2
(NOTE.—This involved membership in a "crowd"—interest in and attachment to others, but without the initiative factor for males, the selection factor for females.)
4. Consistent deep interest in same sex attachments with restricted or no interest in opposite sex 3
5. (a) Casual same sex attachments with inadequate attempts at adjustment to going out with opposite sex 4
(NOTE.—This differs from 4 on the basis of the consistency and meaningfulness of the same sex attachment.)
- (b) Casual contacts with boys and girls 4
(NOTE.—This differs from 3 in that the person was not a regular member of a crowd and just associated with others on occasion.)
6. (a) Casual contacts with same sex, with lack of interest in the opposite sex 5
- (b) Occasional contacts with opposite sex .. 5
7. No desire to be with boys and girls; never went out with opposite sex 6

C. Social Aspects of Recent Sexual Life— 30 Years of Age and Above

1. Married and has children, living as a family unit 0
2. Married and has children but unable to establish or maintain a family home 1
3. Has been married and had children but permanently separated 2
4. (a) Married, but considerable marital discord 3
- (b) Single—has had engagement or deep heterosexual relationship but was emotionally unable to carry it through to marriage 3
5. Single, with short engagements or relationships with the opposite sex which do not appear to have had much emotional depth for both partners, i.e., affairs 4
6. (a) Single, has dated some, but without other indications of a continuous interest in the opposite sex 5
(NOTE.—Implication here is that person has dates every once in awhile but that this behavior is not habitual—

- doesn't play an important part of his/her life (take-or-leave attitude.)
- (b) Single, consistent deep interest in same sex attachments, no interest in opposite sex 5
7. (a) Single, occasional same sex contacts, no interest in opposite sex 6
- (b) Single, interested in neither men nor women 6
- C. (continued) *Social Aspects of Recent Sexual Life—Below 30 Years of Age*
1. Married, living as family unit, with or without children 0
2. (a) Married, with or without children, but unable to establish or maintain a family home 1
- (b) Single, but engaged or in a deep heterosexual relationship (presumably leading toward marriage) 1
3. Single, has had engagement or deep heterosexual relationship but has been emotionally unable to carry it through to marriage 2
4. Single, consistent deep interest in attachments to persons of either sex 3
- (NOTE.—This implies an habitual interest in object relations, a consistent desire for human intimacy, but has never settled into a meaningful, continued relationship with *one* partner in particular.)
5. Single, casual relationships with persons of either sex 4
- (NOTE.—Has dated more often than implied by 6 below, less often than implied by 4 above. Differentiate on the basis of frequency, regularity of social-sexual activity.)
6. Single, has dated a few persons casually, but without other indications of a continuous interest in object relationships 5
- (NOTE.—Dating here the exception rather than the rule. Person has had occasional social-sexual contact, but doesn't actively seek out other persons. This behavior not consistent, nor an important part of his life. His contacts have been solely casual, i.e., with prostitutes to satisfy sex drive; no warmth or capacity to establish human relationships.)
7. (a) Single; never interested in or never associated with either men or women; asocial 6
- (b) Antisocial; destructive, belligerent acting out against others 6

D. *Personal Relations: History*

(NOTE.—Score here is determined by the time of life at which person withdraws,

narrows his range of social contacts. The earlier this occurs, the higher the score will be.)

1. Always has been a leader, and has always had many close friends 0
- (NOTE.—Score for "closeness" if record *states* close friends, or describes frequent contact, shared activity.)
2. Always has had a number of close friends but did not habitually play a leading role 1
- (NOTE.—From childhood until breakdown, person had extensive social contacts.)
3. (a) From adolescence on had a few close friends 3
- (NOTE.—This may involve a drop in the number of close friends after adolescence, but person has retained relationships involving mutual give-and-take with several people through this period.)
- (b) From adolescence on had a few casual friends 3
- (NOTE.—Person maintains relationships with several persons, even though these relationships may lack real emotional depth. Throughout life he has kept up contact with others.)
4. From adolescence on stopped having friends 4
- (NOTE.—Cultivated human relationships during childhood, but has withdrawn since puberty.)
5. (a) No intimate friends after childhood 5
- (NOTE.—Withdrawal began earlier—before puberty.)
- (b) Casual, but *never* any deep, intimate, mutual friendships 5
- (NOTE.—Implies no close friends, even during childhood, but *did* maintain contacts on a superficial level, as distinguished from 6 below.)
6. Never worried about boys or girls; no desire to be with boys and girls 6

E. *Recent Adjustment in Personal Relations*

(NOTE.—Score here the period prior to the noticeable change in behavior which *preceded* symptoms and hospitalization. Any changes noted within 6 months to a year prior to hospitalization will constitute a "change" by this definition. Score period *prior* to these changes.

1. Habitually mixed with others, was usually a leader 0
- (NOTE.—Again, this involves extensive social contacts.)
2. Habitually mixed with others, but not a leader 1

3. Mixed only with a close friend or group of friends	3
(NOTE.—Distinguished from 4 below on the basis of consistency and frequency of contacts.)	
4. No close friends <i>or</i> very few friends <i>or</i> had friends but never quite accepted by them	4
5. Quiet <i>or</i> aloof <i>or</i> seclusive <i>or</i> preferred to be by self	5
6. Antisocial, actively avoided contact, acted out against others	6

Diagnosis and Pattern of Reaction to Drug Treatment: Clinically Derived Formulations¹

*Donald F. Klein, M.D.*²

I. Problem: The Indications for Drug Treatment and the Assessment of Drug Effect

The recent development of relatively effective psychopharmacological agents such as the phenothiazines and imipramine-like compounds confronted the psychiatric profession with the inadequacy of its diagnostic system for the provision of rational indications and contraindications for drug therapy. Within each standard diagnostic group there are wide variations in drug-induced behavioral reactions. Furthermore, there is a marked overlap in response patterns between diagnostic groups, although moderate statistical regularities occur. Because of these limitations, clinical treatment with psychotropic agents has been governed by rule of thumb modified by trial and error. To arrive at generally valid conclusions predictive of drug effect from such clinical empiricism is made difficult by spontaneous remissions, uncontrolled fluctuations of life circumstance,

and the often meager abilities of psychiatric patients to describe their level of cognitive, affective, and social functioning objectively.

The observations and conclusions of this paper are based on experience starting in 1959, with phenothiazines and imipramine, at Hillside Hospital. Initially, we wished to gain a broad image of the patient's behavior and drug effects, avoiding simple lists of traits and symptoms. The use of psychotropic drugs throughout the hospital was restricted, in that they were only prescribed by a research psychiatrist, at the request of the patient's psychiatrist. Prior to starting medication, the patient's case record was reviewed and each patient was interviewed by a research psychiatrist. During drug therapy, the patient's response was assessed through weekly interviews with the patient as well as ward personnel, and through biweekly conferences with the resident and his supervising staff psychiatrist. On each such occasion, the patient's mental status, progress in psychotherapy, and utilization of hospital facilities were discussed.

When it became evident that the standard diagnostic nomenclature was of limited use in categorizing behavioral reactions to drugs, and that psychodynamic formulation lacked predictive clarity, we decided to derive a descriptive behavioral typology of drug effects. In the typological assessment, the detailed records were reviewed by three research psychiatrists, and a consensus statement was made concerning the patient's behavioral reaction during the medication period. The patients were divided into groups on the basis of changes in symptoms, affect, patterns of communication, and participation in psychotherapy and social activity. In each categorization, the behavioral reaction during drug therapy was the determining feature. No attention was paid to the patient's pretreatment behavioral pattern, except as it was relevant to the perceived changes. Inspection of these groups showed, however, that in addi-

¹ Supported, in part, by grant MH-08004 of the National Institute of Mental Health, USPHS. The cooperation of Geigy Pharmaceuticals and Smith, Kline & French Laboratories is gratefully acknowledged. The unfailing collaboration of the staff of Hillside Hospital made these studies possible.

² Director of Research, Hillside Hospital, Glen Oaks, N.Y.

tion to the similar behavioral changes with drug therapy, their members had similar developmental and behavioral pretreatment characteristics.

These initial studies of the phenothiazines and imipramine were conducted on approximately 200 patients. The clinical patterns of drug response reported have continued to be substantiated by further experience (1, 2).

Following the clinical phase an experimental program was initiated in 1961 to study the effects of chlorpromazine and imipramine. All psychotropic drug therapy at Hillside Hospital was initiated by referral by the patient's supervising staff psychiatrist to the research team. Pretreatment examinations consisted of research psychiatric interviews, behavioral ratings by the patient's psychiatrist and ward personnel, self-ratings of symptoms and attitudes, psychological tests, EEG, medical examination, and physiological tests, including blood pressure response to mecholyl and radioactive iodine uptake (3-9).

Subjects were randomly assigned to placebo or one of two fixed dosage increment drug schedules for 4 weeks: imipramine, 75, 150, 225, and 300 mg.; chlorpromazine, 300, 600, 900, and 1,200 mg. Each dose of chlorpromazine was combined with proportionate amounts of procyclidine, 3.75, 7.5, 11.25, and 15 mg., to prevent akinesia and other extrapyramidal disorders. Each medication was dissolved in a highly flavored liquid placebo vehicle and each patient received 40 ml. per day from individually labeled bottles. Maximum dosage was maintained for 2 weeks and retesting was conducted during the sixth week of medication.

One hundred seventy-three patients started medication and 150 completed the study program. During the 7 weeks of the experimental program, the longitudinal clinical observations of the pilot phase were maintained. In addition, following the experimental period all patients were fol-

lowed throughout their hospital stay so that their clinical response to other interventions could be assessed. To gain information concerning the long-term outcome, both for diagnostic clarification and estimates of treatment efficacy, all patients entered into this program were followed up by detailed interviews of themselves and informed relatives 2 and 3 years after hospital discharge (10).

It should be emphasized that all referrals for psychotropic medication were accepted regardless of diagnosis or symptomatology, since it was hoped to be able to utilize the drug effects as tools for discriminating subgroups of patients thus requiring a large heterogeneous group. Such division of the drug groups led to small patient response subgroupings. Therefore, in an effort to expand our data the experiment was replicated, starting in 1963. This second experiment is now being terminated after an additional 150 patients completed the study.

For this presentation, the primary emphasis is on the detailed clinical psychiatric categorization of the patients in terms of drug response with its diagnostic and psychopharmacological implications. Also, certain changes in the usual study methods are described as necessary for the prediction of drug effect.

II. Clinical Groupings and Theoretical Inferences

A. Ordering Effect of Drugs on Psychiatric Data

Certain aspects of drug treatment deserve emphasis as being peculiarly useful in the attempt to understand psychiatric patients. Each patient presents a welter of unique familial relationships, developmental idiosyncracies, social aberrations, affective states, cognitive abilities, symptoms, defects, and maladaptations. The information overload upon the diagnostician is so acute that he is often led to premature closure and scotomization of the data. Further-

more, since the clinician is usually confronted with a cross-sectional picture it is very difficult to know which manifest pathological features are primary, in the sense of being tightly linked directly to the fundamental disturbances, and which are secondary in the sense of being inconstant accompaniments or sequential reverberations.

As Ledley and Lusted point out, the fundamental formula of medical diagnosis is "if medical knowledge E is known, then if the patient presents symptoms G , he had diseases f ." Conversely, "if diseases f are cured then the patients' symptoms will disappear" (11). Unfortunately for the neatness of this formulation, it does not distinguish reversible symptoms from irreversible effects. The persistence of painful, ineffective, and maladaptive affects, thoughts, beliefs, and behaviors in psychiatric patients, long past the period when the apparently initiating causes have ceased to operate, is one of the core problems in pathogenetic theory. If this persistence is due to a variety of causes, rather than some unitary repetition compulsion, it is possible that the identification of drug effects which terminate some persisting difficulties, ameliorate others, and are ineffective in still others may aid us in approaching this aspect of pathogenesis systematically. "Curative" drug therapy would allow us to determine which of the manifold aspects of the patient are symptoms of the illness, and which aspects are illness irrelevant.

B. Psychiatric States With Complete or Almost Complete Remission

Attention to this question of complete symptomatic remission was attracted by the frequent remarkable psychopharmacological successes in certain groups; i.e., retarded depressions, manic states, and angry hyperactive paranoid states.

1. Mood elevations in retarded depressions.—These patients are depressed, inactive, self-depreciatory, hopeless, and un-

interested in the environment. They relate in an apathetic, uncommunicative, and passive fashion and complain of the sudden onset of anorexia, insomnia, unhappiness, loss of interest in their activities, inability to cope with their responsibilities, to concentrate or make decisions. With imipramine treatment there is a distinct increase in environmental interaction and goal directed activity followed by a lessening of complaints and a general cheerfulness of mood. Complaints of tension and apprehension related to difficulties in discharge planning and returning to the community remain, however, with anxious procrastinating attempts to evolve a foolproof plan.

2. Reduction of anger and psychomotor acceleration in manic and angry, hyperactive, paranoid patients.—One group of patients caused considerable diagnostic disagreement. Usually they were labeled manic-depressive, manic- or schizo-affective, or schizophrenic, paranoid type.

These patients approach the staff in an angry, supercilious, and manipulative manner. They are demanding of their doctors and are frequently suspicious and evasive when pressed concerning their feelings and personal relationships. There are frequent complaints of unjust confinement through the blundering errors of friends, relatives, and hospital staff. They express their anger elatedly, without marked fearfulness or depression, often revealing persecutory delusions, ideas of influence, sexual abuse, and auditory hallucinations. Flight of ideas, loss of goal orientation, and illogical associations are demonstrated. On the ward they demand attention, are meddlesome and verbally aggressive. They are generally too hyperactive and negativistic to engage in hospital activities.

The response to medication of this group is a reduction of anger, psychomotor deceleration, and increased social participation. These patients require large doses of chlorpromazine (about 2,000 mgs. per day) for effective control. This dosage is easily

tolerated when given in combination with anti-parkinson medication.

Patients now approach the interviewer in an ingratiating manner, expressing the understanding that they had reacted in an emotionally disturbed way in the past. The greater their previous delusional formation, the greater their denial of illness. There is a decrease in expressions of angry frustration and an associated increase in tolerance for frustration. Psychomotor acceleration disappears and mood is more evenly maintained, with infrequent mild depressive dips. In psychotherapeutic sessions they are capable of more meaningful discussions of their interpersonal relationships and feelings.

They interact more freely with both patients and staff in a helpful but often controlling manner. They express boredom with hospital routine and contrast it unfavorably with the full and active life they would be leading at home. The staff evaluated these patients as greatly improved, and occasionally used such descriptions as "cured" or "remitted." Prognosis was also rated "hopeful."

Under drug treatment each of these groups frequently shows a complete or almost complete restitution to the pre-morbid status. The fact that the natural history of each of these conditions is also marked by periods of apparently complete spontaneous remission is a striking parallel that can hardly be an accident. The outstanding commonalities of these illnesses is their relatively late onset in adult patients with good premorbid social and psychological development, and the marked derangement of psychomotor activation and mood.

The fact that both drug treatment and the natural course of these diseases are marked by complete remissions lead me to postulate that the core of these illnesses lies in their common denominator, a reversible pathological mood and activation state, and that both the processes of natural repair and drug treatment operate by normalizing this central pathological state.

Cognitive defects that occur in these states seem to be either related to the state of psychomotor acceleration or mood congruent convictions. That is, both the manic and the angry, hyperactive paranoid show distractability, flight of ideas, impulsiveness, and loss of goal organization. The retarded depressive, on the other hand, has paucity of ideas, tenacious ruminations, and inability to convert ideas into action spontaneously. Delusions are mood congruent; e.g., the belief in unworthiness and physical illness in the depressives, the conviction of grandiose abilities in the manic, or the certainty of the hostility of psychiatric staff and relatives in the angry paranoids.

Both the mood disturbance and the psychomotor rate disorder appear to ameliorate simultaneously. The convictions, which initially may be attempts to conceptualize and rationalize the subjectively overwhelming, radically altered mood state, have a degree of functional autonomy persisting past the termination of the originally causative mood and cognitive disturbances. Therefore, the depressive may be left with lowered self-esteem, the manic may retain a conviction of uniqueness, special ability and creativeness, and the paranoid may remain somewhat suspicious and wary of an environment seen as potentially hostile.

C. Psychiatric States With Functionally Autonomous Residual Pathology

1. Reduction of episodic anxiety in phobic-anxious states.—This schema of patients with primary reversible affective-activation disturbance and secondary affective reactions, and conceptual convictions which have relative functional autonomy received further support from the study of phobic-anxious states (12).

These patients also had episodic illness, in that their lives were punctuated by discrete series of overwhelming attacks of panic. Typically, subjects note the sudden onset of inexplicable "panic" attacks, accompanied by rapid breathing, palpitations, weakness, a feeling of impending

death, and occasionally depersonalization. Their activities become progressively constricted, until they are no longer able to travel alone for fear of being suddenly rendered helpless while isolated from help. Depressive complaints are frequent and associated with feelings of futility. Although fear of open spaces is not the hallmark of this condition but rather expectant fear of lack of support when overwhelmed, their condition is often referred to as agoraphobia. They engage in prolonged outpatient psychotherapy, usually devoted to the exploration of unconscious sexual and aggressive impulses, with the interpretation of the phobically barred areas as situations of forbidden symbolic temptation. Hospitalization occurs after the family can no longer tolerate the restrictions placed upon it.

Under imipramine treatment the "panic" attacks cease, although both phenothiazines and sedatives had been previously ineffective. However, the patients are reluctant to change their phobic behavior pattern, having shown the same reluctance following spontaneous terminations of panic attacks. Because of the persistence of this secondary apprehensiveness, the patient may complain and behave in an unchanged fashion after imipramine treatment has terminated the panic attacks. These patients require supportive and directive treatment focussed on changing the conceptual residue of the original disturbance (i.e., "I might be overwhelmed by panic, at any moment"), for manifest behavioral change to occur. However, such supportive and directive treatment is entirely ineffective in the face of continued panic attacks.

Imipramine has been an important aid in the definition of this nosological group since these patients also present a wide array of associated obsessional, hypochondriacal, affective, dependent, passive-aggressive, addictive, histrionic, and manipulative features. Depending on the salient symptomatology, or the selective perception of the diagnostician, they may be referred to

as obsessionals, hysterics, atypical depressives, passive-aggressive character disorders, pseudo-neurotic schizophrenics, acute schizophrenics, alcoholics, barbiturate addicts, anxiety states, conversion reactions, etc. This diagnostic chaos is the natural result of the psychiatrist's inability to arrive at a hierarchy of importance for this multiform symptomatology on any cross-sectional basis. Attempting to arrive at a consensus between dissenting diagnosticians usually results in a fence-straddling label; e.g., severe mixed psychoneurosis or pseudoneurotic schizophrenia or borderline state, etc.

Imipramine treatment highlights the panic attack as the proximal cause that both initiates the adaptive dependent-phobic behavioral sequence and maintains it by an irregular reinforcement schedule. Thus, it gives a marked positive weighting to such a symptom in the attempt to predict drug effect and to form homogeneous diagnostic subgroups. This is further emphasized by the fact that phenothiazines, which are often considered antianxiety agents, are positively deleterious in their effects on this syndrome although they are most helpful in "panic" episodes that are part of acute psychotic episodes marked by delusions hallucinations, or referential thought disorder. This group furnishes a good example of the utility of patterned predictors in the development of drug-relevant diagnostic categories.

2. *Reduction of recurrent agitated depression in obsessive-compulsive states.*—This schema of primary drug responsive activation-affective disturbances and residual affective or cognitive symptoms which are functionally autonomous led to attempts to see if other disorders approximated this sequence.

Certain patients who are often diagnosed as "Agitated Depressions in Obsessive-Compulsive Personalities" may be hospitalized because of the extent of their inability to function and the manifest risk of suicide. Their self-presentations are dominated by descriptions of battles against ego-

alien obsessions and compulsions. Their emotional states vary from a chronic discouragement through fearful ingratiation in the hope of help, to ludicrous outbursts of coercive agitation coupled with threats of suicide and demands for immediate succor. Their communication is well organized and no thought disorder is evident outside of the obsessional preoccupations. Psychotherapy is barren, repetitive, ruminative, intellectualized, and symptom-centered. On the ward, the patients are ritualistic and solitary but tractable, except during their demonstrations of despair and hopelessness. Insomnia is common.

Several alternative formulations concerning this cross-sectional symptom state are possible. First, the agitated depression may be a secondary psychological despair reaction to the primary life-crippling obsessions and compulsions. Another formulation might be that the obsessions and compulsions are secondary, magical, self-punitive, expiatory attempts to alleviate the primary depressions. Still another formulation might be that the obsessions, compulsions, and agitated depression are all secondary reactions to, or attempts to control, a primary underlying disturbance, e.g., breakthrough of unconscious hostility.

Phenothiazine treatment in such cases induces the following sequence of changes. The agitation and insomnia are sharply reduced within a few days so that ward personnel and family are immediately relieved by the change. The depressive complaints and lack of social participation changes more slowly, but with a stimulating environment the patients participate more actively within a few weeks and finally state they are no longer depressed. The obsessions and compulsions hardly change at all in frequency although the patients more easily disregard these symptoms (2).

Again study of the natural history of the illness shows a striking parallelism with its clinical course in response to drugs. These patients were obedient, fearful, dependent children who developed lifelong ob-

sessional traits. Later, episodic agitated depressions resulted in acute failures in psychosocial functioning followed by rather lengthy convalescences, with prolonged exacerbations of obsessive-compulsive symptoms even after the patient returned to social functioning. Drug treatment restores them to their typical obsessional mode of functioning prior to the present episode of depressive decompensation, in a fashion similar to spontaneous remission.

Both on historical and drug effect grounds, we cannot maintain that obsessions and compulsions cause depression, or that depression is necessary for obsessions and compulsions, or that both are conjointly caused by a third present factor, since the depressive state fluctuates independently of the chronic obsessive-compulsive state. Perhaps both obsessive-compulsive fixed symptomatology and recurrent agitated depressions are common enough so that they may occasionally randomly co-exist in one person. Still other data would be necessary to resolve this issue; i.e., a community survey giving the incidence of these disorders separately and jointly so as to determine if their conjoint presence is simply due to random co-existence. However, I believe most psychiatrists would hypothesize a high incidence of agitated depression in obsessional states. Both obsessive-compulsive symptomatology and recurrent agitated depressions may reflect a common early psychological and constitutional difficulty involving separation anxiety, proneness to depression, marked dependency, and severe ritualistic conscience formation. This difficulty has eventuated in both a functionally autonomous, early developed, chronic, drug-unresponsive, obsessive-compulsive symptom pattern in association with a persisting drug-responsive vulnerability to depressive disorder.

3. *Affective stability reactions in emotionally unstable character disorders.*—Another group of patients who approximate an affective-activation disturbance with a functionally autonomous residue are often

diagnosed as "Emotionally Unstable Character Disorders," although underlying or incipient schizophrenia is usually considered. The affective-activation disorder of these predominantly adolescent patients consists of alternating short periods of tense, empty dysphoria accompanied by inactivity and withdrawal, suddenly shifting to periods of impulsiveness, giddiness, low frustration tolerance, intolerance of rules, and short-sighted hedonism. This marked lability is often not immediately noted as core pathology because of the complicated self-presentations which range from a fragile, immature, dependent image, eliciting protectiveness from the observer, to a hard "wise guy" presentation appearing as independence and lack of need for support. The patients are perplexed about their goals in life, stating that they do not know who they are, what they are, or what they want to be. They are confused over issues of dependency, intimacy, and self-assertion, reacting in a disorganized, flighty, and despairing fashion. There is a pervasive feeling of being excluded from normal life and from peer groups, with a conviction of being irreparably bad. They are rational, relevant, and coherent except for periods of either giddiness or agitated fearfulness during which their speech and behavior become scattered and disorganized. Participation in hospital activities fluctuates considerably, but when involved, these patients are often creative, skilled, and original.

With phenothiazine treatment there is a decrease in affective lability and impulsiveness. A bland, placid, friendly, and ingratiating manner replaces feelings of confusion, perplexity, anxiety, and depression. Feelings of role diffusion and inability to find a path in life decrease without finding specific solutions. Flight of ideas is no longer apparent, and during psychotherapy their introspective ruminations are changed to a concern for day-to-day events and for peer relationships, without interest in long-range planning. The patients rejoin their peers in a friendly fashion and

are active and popular. In place of the rapidly alternating phases of activity and withdrawal there is a more uniform degree of friendly interaction, but occasional shortsighted impulsive actions may still occur.

Thus, with phenothiazine medication their lability decreases, their frustration tolerance increases, and their social behavior becomes more acceptable. However, the capacity to think of long-range goals and step-by-step achievements remains absent. They experience little distress in their present lives and are unable to view the future meaningfully. The stage is thus set for a bland denial of difficulty coupled with an indifference to productive work and planning, and occasional impulsive hedonism (2).

In these cases the secondary residue consists of a mixture of sociopathic attitudes and planning defects that seem to be partly the consequence of their previous inability to rely upon themselves as a stable base, combined with their previous experience of a kaleidoscopic world. Also during phenothiazine treatment, they are particularly prone to complain of a sense of boredom, deadness, and inability to really enjoy themselves, reminiscent of the opiate addict's complaint of feeling badly when "on the natural." Their complaints of pervasive anhedonia seem not to be borne out by social observation since they frequently may be observed evidently enjoying themselves and are highly motivated to organize parties, outings, etc., that may provide them with "kicks." One may hypothesize that this complaint, which occurred prior to drug treatment, may be due to contrast with their previous episodes of gleeful elation. In a sense they are addicted to these "high" periods, miss them, and are willing to suffer the discomforts and disruptions of mood lability to regain these affective states. Therefore, another functionally autonomous residue consists of the memory of pathologically pleasant states with the contrast derogation of normal mood.

Phenothiazine treatment highlights the disruptive effect of the short-term mood swings as central to this group's psychosocial decompensation. These patients are often capable of supervised work and recreation, while receiving medication, in spite of the persistence of their social attitudes. Again there is a parallelism of drug effect and natural history. Although insufficient longitudinal data has been gathered to be definitive, these patients may have a relatively good prognosis, their lability, and impulsive antisociality "burning out" with increasing age, leaving a socially acceptable citizen with somewhat deviant manipulative and irresponsible attitudes.

A patient group which resembles the "high" phase of the "Emotionally Unstable Character Disorders," are chronically ebullient, overactive, impulsive, rash, with marked lack of foresight, and intolerance of routine and rules. The qualitatively abnormal attitudes and social maneuvers are figurally salient over the background activation-affective state so that they may be diagnosed as "Passive-Aggressive Character Disorders, Aggressive Type." This diagnosis fails to recognize the chronic activation dysregulation, whereas the European term "Hyperthymic Psychopath" seems more appropriate. After phenothiazine treatment these patients are indistinguishable from the drug-treated Emotionally Unstable Character Disorders.

4. *Reduction of schizophrenic disorganization.*—The relationship of the manic-depressive states to the schizophrenias has been a perplexing question since the original distinction was made by Kraepelin on the basis of course: Schizophrenics deteriorating and manic-depressives remitting periodically. However, the existence of remitting patients with unquestionable episodes of thought disorder caused the development of many diagnostic labels, including that of schizo-affective psychosis, which pointed to the lack of a simple correlation between disordered thinking and deteriorative course.

It is generally acknowledged that prognosis in schizophrenia, meaning remission with the relative absence of persisting deleterious residual defects, is best in patients with familial history of psychotic depression, acute adult onset, precipitating cause, affective turmoil, depressive preoccupations, confusion, good premorbid personality, and socially adequate adjustment. In other words, the prognosis is better the closer the onset, heredity, and symptomatology approximates manic-depressive illness, and to a lesser degree, toxic states. Again, prognostic criteria for phenothiazine treatment of schizophrenia are similar to those for prognosis in the natural course of the disease in that those patients with acute onset, affective turmoil, good premorbid personality, and manic-depressive features respond best, with least residual defect.

The fact that both acute manic and acute schizophrenic episodes are produced by imipramine in psychotics and patients with CNS damage but not in normals, neurotics, or character disorders (13), may indicate, on the grounds of parsimony, that these naturally occurring acute psychotic states share a common pathophysiological link that is sensitive to the effects of imipramine. It is often very difficult to distinguish acute manic from acute schizophrenic states and the diagnosis is frequently made on the basis of the amount of residual cognitive and affective disorder remaining after the acute flareup. The common ameliorative effects of phenothiazines in these conditions may also indicate action at a common pathological site.

One might attempt to integrate these pharmacological and nosological considerations concerning schizophrenia and manic-depressive disease by the following considerations:

(a) The common pathophysiological link in the acute psychotic state consists of a disorder in the interconnected normal neurohumoral mechanisms regulating arousal, activity, pleasure response, fear,

and anger. The existence of such "activation" mechanisms has received convergent support from many sources, reviewed in Duffy's extensive monograph (14). The activation mechanisms seem to have evolved in relation to states of heightened environmental interaction and organismal activity as exemplified in waking, flight, fight, and pleasurable activity. Obviously, all of these states are crucial for survival.

Duffy points out that behavioral description seems to require regularly the use of two descriptive categories: (1) "direction" (relating to selective environmental evaluation leading to approach or withdrawal) and (2) "activation" (relating to intensity of response tending towards overt expression). "Activation" should not be directly equated with overt activity because of the possibility of inhibitory intervention. "Emotion" may be considered as a compound phenomenon of "activation" and "direction," although the issue of arousal patterns specific to fear and anger remains debatable. The fact that two independent descriptive categories are necessary leads to the inference of two partially independent systems of behavioral regulation, each having specific pathological deviations.

(b) It is postulated that only the range of emotionally labile, manic, and depressive disorders will be produced if only activating mechanisms are pathologically involved. Activation dysregulation may occur in two ways. Through the postulated pathological disturbance in activation regulation or through states of extreme environmental stress producing excessively intense and prolonged activation.

(c) In some schizophrenic conditions it is postulated that there are other forms and degrees of central nervous system pathology involving cognition and selective evaluation of the environment, ranging from the severe to the subclinical, exclusive of pathology in the activating systems.

Pollack (15) has pointed out that there is a significant relationship between psychiatric patients' I.Q. and age of hospitali-

zation. The lower the age of hospitalization (which in turn is positively correlated with the age of appearance of manifest illness), the lower the I.Q. Discharge ratings are correlated both with age and I.Q., the poorest occurring in young patients with low I.Q. This is compatible with a range of severity in CNS "directive" defects in schizophrenics, the most severe defects causing the most obvious and therefore early detected behavioral disturbances. These CNS defects induce a range of chronic disorder in the directive, integrative, cognitive functions and become more disruptive during periods of activation dysregulation.

It is postulated that in other schizophrenics, pathology in the activation system induces an endogenous toxin that causes interference with the directing system and that in these patients there is no initial defect in the directing system. Relevant to this is the finding that the plasma factor in schizophrenia is probably not present at all times and that its level is closely related to the physical activity of the subject.

(d) This schema would predict a two-stage process in the pathophysiology of some acute schizophrenias characterized by sequential pathology in both activating and directing systems. An initial stage of mood disturbance and insomnia, is superseded by angry hyperactivity, psychomotor acceleration, and flight of ideas shared by acute mania, or loss of environmental interest shared by acute depression, followed by gross psychotic perceptual and cognitive disruption, including delusions and hallucinations. This sequence has been observed by Klein and Fink (6), Pollack et al. (5) using imipramine, and Brune and Himwich (16) who used MAO inhibitors in the treatment of schizophrenics. These groups reported a biphasic psychogenesis: a period of mood elevation approaching inappropriate euphoria and increased environmental interest, followed by angry hyperactivity and cognitive disorganization.

(e) The prognosis may be best in schizophrenic patients whose onset and symptomatology approximate manic-depressive disease, simply because they required the greatest activation dysregulation to produce cognitive disorganization. The implication is that manic-like schizophrenics have the least CNS cognitive defect and that following the period of affective turmoil they are closest to normal CNS functioning, and therefore best able to manage their lives in a fashion approximating the norm. The prognosis may be worst in schizophrenics with early, insidious onset and flattened, inappropriate affect, because the onset of cognitive disorder in the absence of activation disorder indicates a more profound CNS defect that is less likely to become compensated.

(f) The process-reactive continuum hypothesis of the schizophrenias resembles this model: process schizophrenics may be equated to patients with major CNS cognitive defects and reactive schizophrenics to those with minor CNS cognitive defects activated by a pathological arousal mechanism or an intact directing system deranged by an activation-linked toxin. It should be noted that this interpretation is at variance with the usual one that equates process-reactive with organic-psychogenic. It is generally recognized that for some remitting schizophrenias each psychotic episode is followed by a decrement in the patient's affective and cognitive level, leading to chronic defect after repeated episodes. In these cases the residual symptomatology may be a mixture of irreversible CNS defect and a functionally autonomous mixture of self-protective maneuvers such as social self-isolation, passive-dependent relationships, unwillingness to engage in problem solving activity, etc.

Again the natural course of illness and the results of drug treatment seem isomorphic, since those schizophrenic patients who most closely resemble manic-depressives respond best to phenothiazines, with mini-

mum residual symptomatology. Therefore, one might predict that the new phenothiazine treatment of schizophrenia should result in no increase in the percentage of patients who are symptom free at followup since this would be fixed by the natural incidence of manic-depressive-like schizophrenics. However, there should be a shift in prevalence from those patients who are grossly psychotic to those who may be maintained in the community with discernible residual defect. Kelly and Sargant have reported such findings (17).

D. Drug-Refractory States

Just as patients with complete remission highlight the common denominators of drug action, drug-refractory patients highlight drug irrelevant motivational, affective, and cognitive states.

1. Patients without manifest activation-affective dysregulation.—It follows from the theory that psychotropic agents work reparatively on activation-affective dysregulation that conditions not characterized by this defect should be refractory to these drugs. On the whole this is accurate. Dys-social, antisocial, passive-aggressive, and narcissistic character disorders, perversions, obsessive-compulsive states, mental deficiency, conversion and hypochondriacal reactions, chronic "burned out," and pseudoneurotic schizophrenics are all refractory to phenothiazines and imipramine-like drugs. Whenever drugs have proved useful in these conditions it is in the treatment of an intercurrent affective derangement, without alteration of the patients' chronic difficulty. Again drug treatment mimics natural course; i.e., unremitting chronic symptoms with transient affective flare-ups.

2. Patients with apparent activation-affective dysregulation.—(a) Patients whose specific defects elude drug treatment. It must be admitted that patients who conform exactly to the descriptions given above of affective-activation disorder may not respond to psychotropic agents. Under the

usual clinical circumstances it is difficult to try one antidepressant or phenothiazine after another, methodically, in search of a suitable agent, especially since ECT is regularly effective in these drug-refractory states. However, certain depressives who are refractory to imipramine respond to desmethylinipramine or an MAO inhibitor. It seems likely that these therapeutic specificities are due to different specific defects in the basic pathophysiology that are not reflected in the manifest symptomatology. Careful comparative studies of diagnostically and phenomenally similar patients who respond differentially to drugs may help isolate such basic pathophysiological differences.

(b) The histrionic. One patient group who causes great difficulty for any theory of psychotropic drug action, as well as grave problems in clinical management, is comprised of characteristically labile, episodically agitated, erratic, unpredictable, manipulative, tense, and histrionic patients. They are usually rational, relevant, and coherent, although there are occasional paranoid and hallucinatory verbalizations. They may express themselves as being panicky and frightened to the point of suicide and several minutes later be affably laughing with other patients. Their therapists readily become emotionally involved and are frequently frustrated and perplexed by the inability to predict or modify the patient's behavior. The patients express great investment in their doctors and psychotherapy, endowing them with miraculous potentialities. They maintain a high degree of interaction with other patients, being at times sociable, friendly, and supportive, and at other times disturbing, hostile, argumentative, and demanding, and yet again pleading for help and direction.

There is no sustained response to medication, although a transient initial improvement is common. Intolerance of distressing side effects is expressed dramatically. Because of their lability, there is much uncer-

tainty as to the effectiveness of medication, so that they tend to receive long treatment courses before the medication is found to be ineffective by the staff. However, if the patient responds in a somatizing fashion medication is terminated promptly.

Why this emotionally labile and agitated group should prove so refractory to phenothiazine medication remains a real problem. One clinical feature of predictive value and theoretical import is the marked relationship of their symptoms to environmental impact. Unlike depressed patients who may increase the vigor of their complaints in the presence of psychiatric staff in an attempt to coerce a maximum curative effort from them, but who remain inactive or unproductively agitated when not under staff observation, this group of refractory patients may appear in good spirits, when apparently unobserved by staff, engage in social games and gossip pleasantly with other patients, even shortly after an explosive affective display. The key issue may be histrionic role playing and symptom imitation, in other words, that these patients' symptoms may not be the direct external manifestations of intolerable affective states, but rather may be environmentally oriented, learned, manipulative devices. Agitation and depressive complaints in this group are not the same as similarly labeled phenomena in other patients.

Psychiatrists often comment on the disappearance of the classical hysteric; i.e., hysterical fits, paralyses, and anesthetics are conspicuously absent on our psychiatric wards. In Charcot's time the practice was to hospitalize hysterical patients with epileptic and neurological patients. It seems likely that hysterical symptomatology was modified by iatrogenic exposure. Nowadays hysterics (manipulative, histrionic, imitative, labile patients) are hospitalized with schizophrenics and it may be that we are thus producing what might be termed "pseudoschizophrenic neurotics" or "hyster-

oschizophrenics" in the same fashion as hystero-epilepsy was produced 80 years ago. These patients may supply a considerable proportion of the surprising aberrant somatizing and negative responses to phenothiazines on the part of apparently schizophrenic patients.

(c) The chronically anxious. Another group of patients who are comparatively refractory to phenothiazines or imipramine are those with chronic anticipatory anxiety. A pharmaco-diagnostic note is the moderate usefulness in these patients of drugs of the sedative class; i.e., alcohol, barbiturates, meprobamate, glutethimide, ethchlorvinyl, chlordiazepoxide group. The usefulness of these agents is limited by pharmacological tolerance since increasingly larger doses are required to maintain equanimity, and these patients often self-administer these compounds to the point of addiction. This is in sharp contrast to the effects of the phenothiazines and imipramine where progressively less medication is necessary to maintain remission following initial effective treatment. Also in sharp contrast is the general uselessness of these sedative agents in conditions of marked activation-affective dysregulation.

Some theories of phenothiazine effect place heavy emphasis on the central role of antianxiety actions. That is, various symptoms such as hallucinations, delusions, and agitation are seen as secondary responses to anxiety, and the ameliorative effects of the drug are ascribed to reduction of anxiety. This type of theory has been supported by the reports of supposed specific effects of phenothiazines on conditioned avoidance reactions. The clinical refractoriness of chronically anxious patients to psychotropic agents casts doubt on this.

This reinforces the necessity for discriminating among anticipatory anxiety, panic anxiety, anxiety accompanying affective dysregulation and anxiety in acute schizophrenic psychosis.

III. Implications for Psychopharmacological and Diagnostic Research

A. Diagnosis-Irrelevant Target Symptoms as Predictors of Drug Effect

The poor fit between standard diagnosis and prediction of drug effect led to several attempts at solutions outside the framework of diagnosis. Describing certain invariant symptom-drug relationships, e.g., reduction of agitation, or reduction of anxiety, or mood elevation, was attempted. However, the clinical experience detailed above showed that psychopathological phenomena, often referred to as target symptoms, that were labeled identically in different patients might respond differently to the same medication. Irregularities in the drug response of phenomenally similar symptoms cast considerable doubt on the legitimacy of classifying behaviors together on the basis of gross behavioral or introspective report similarities.

Another approach is to see each symptom as a prominent aspect of an integrated disease that cannot be abstracted from the symptom complex without leading to false equivalences with similar aspects of entirely different configurations. However, this viewpoint requires that diagnostic formulation be systematic patterned multivariate descriptions. Patterned clinical diagnosis is not the same as the products of many of the currently employed multivariate techniques such as person correlations, linear additive multiple regression analysis, or canonical correlations.

For instance, when the trait of "panic" is associated with auditory hallucinations, it should be weighted positively for beneficial chlorpromazine effect. However, in the presence of agoraphobia it should be weighted negatively (1, 2, 12). The implication is that "panic₍₁₎" is simply not the same as "panic₍₂₎" and identical coding of these two phenomenally similar behaviors does violence to the possibility of determining their relationships correctly. Formulating such complexities requires

interaction terms that are not part of the usual additive multivariate approach. Furthermore, these interaction terms are very difficult if not impossible to establish on a statistical inductive basis. They may require stipulation derived from clinical experience. In turn this would require a new level of creative interaction between clinicians and statisticians.

Once patterned clinical diagnoses were developed it would still be necessary to show that the application of these patterns led to predictions of drug effects on specific aspects of the pattern that would be significantly better than predictions that utilized data in univariate fashion or in the form of linear additive multiple regression equations. In any case, the issue of diagnosis has not been avoided by the attempts to relate symptom reduction directly to drug effects, but rather has been reemphasized as an investigative area.

B. Inferred Underlying Universal Variables as Sites of Drug Action

The need to diagnose so as to predict drug effect might be avoided if drug effects are mediated by certain diagnosis-irrelevant underlying variables that are either increased or decreased in their functional capacities by drugs; e.g., depressing the cortex, releasing the subcortex, depressing or stimulating the hypothalamus or reticular activating system, etc. This postulation form is not restricted to neurophysiological hypotheses. One might postulate changes in excitatory or inhibitory processes in Pavlovian or Hullian styles, or decrease or increase in repression or psychic energy in psychoanalytic terminology.

An analogy is with an overly loud radio. The specific fault may lie in many places but simply decreasing the aerial length would provide sufficiently decreased input to compensate for many of them. The actual defects would not be specifically remedied or diagnosed. Aerial shortening may be considered analagous to symptomatic compensatory treatment. One clin-

ical example might be aspirin, whose analgesic effect seems independent of the specific source of painful afferent input. If psychotropic drug effects follow a nonspecific compensation model, their relevance to the elucidation of specific defects is dubious.

This approach presents the same difficulties as the target-symptom approach, with the added problem of getting agreement as to the existence, or granted this, the state of the inferential underlying variable. Also, this model of unidirectional functional effect is hard to reconcile with the normalizing effects of psychotropic agents that frequently ameliorate both extremes of paired polar systems, such as mania and depression, stupor and agitation, fearful withdrawal and assaultiveness.

C. Inferred Pathological Disturbances as Sites of Drug Action

Still another alternative view is that psychotropic drugs specifically affect pathological disturbances in a normalizing reparative manner rather than a compensatory one. In a previous paper I have presented evidence that the pertinent psychiatrically relevant actions of psychotropic drugs are manifest only in pathological cases and present a cybernetic theory of the nature of the relevant pathology and treatment effect (13). Briefly, the phenothiazines and imipramine-like agents are seen as acting by repair of an impaired control mechanism, a pathologically insensitive level detector in an activation-affect controlling servomechanism. This hypothesis has the merit of accounting for the normalizing effect of these drugs, in apparently opposite pathological states, by one pharmacological action. In this view, drug effect and diagnosis are closely tied to each other, in that ameliorative effect implies a specific defect, although the specific cybernetic defect hypothesized may well be incorrect.

These views of drug effect, i.e., compensatory and reparative, are not mutually

exclusive. It is possible that specific physiological drug effects are compensatory for certain pathological conditions, reparative for others, and irrelevant or toxic for still others. There is some heuristic value in this emphasis on the possibility of reparative drug action since it fosters the use of clinical drug effects as the basis for a drug-relevant psychiatric typology. That is, patients would be grouped on the basis of pattern of drug-induced changes in symptoms, affect, patterns of communication, participation in psychotherapy, and social activity. The responses to medication would be used as dissecting tools to distinguish various subpopulations and to permit the determination of specific developmental, physiological, and psychosocial commonalities within each subpopulation. Hopefully these commonalities may shed light on the question of the etiology and pathogenesis of psychiatric disorder, as well as serve the practical purpose of providing rational indications and contraindications for drug therapy.

D. Longitudinal Analysis

One aspect of diagnosis which is often insufficiently attended to is the importance of the extended time sequence of pathological manifestations. Diagnosis in general medicine is never restricted to cross-sectional evaluation but depends heavily on history. Most computer attempts at diagnosis have utilized either cross-sectional examinational material (18) or have used historical data (e.g., age of onset) as entries in a multiple-regression analysis that cannot use information derived from longitudinal ordering. Although change scores may be used as items, extended sequences cannot be put into this form. This issue has probably not been handled adequately because of the unmanageably large number of equiprobable sequences generated by a combinatorial approach.

Also sequential order may be more important than the exact timing of phases. With a form that samples the patient's his-

tory at fixed intervals, the described order of the phases will vary depending on the synchrony of the sampling period with disease periodicity. One might have a patient who was functioning relatively normally until he had a three-month period of emotional lability, followed by a one month period of elation, followed by a three month period of depression, perplexity and withdrawal, culminating in a referential and delusional psychotic state. If one were to use a fixed time-sampling period (e.g., a patient description every 6 months) quite different orders (e.g., normal depressed psychotic, normal manic psychotic, normal labile depressed psychotic, etc.) might be generated depending on the date taken as the zero point. Such fixed time-sampling approaches are less appropriate the more one is attempting to capture a rapid flux, requiring such short intervals as to approach a narrative, and of least interest when one is attempting to describe a slow-moving, nonperiodic condition.

One approach to this problem would be to formalize the descriptive psychiatric knowledge of disease courses in the form of modal sequential descriptions. The course of each patient could be rated for the overall degree it approximates known longitudinal patterns with allowance made for missed phases, and attention paid to qualitatively unique aspects and sequences. Discrepancies would be noted and cases that did not approximate known patterns would be investigated for new patterns. This diagnostic program would emphasize an interplay between the substantiation or contradiction of past beliefs and new inductive attempts rather than a *de novo* inductive attempt.

The relevance of this discussion of the methodology of handling ordered sequential data to the study of the relationship of drug effects to diagnosis is that the problem of analyzing history is formally identical to the problem of analyzing the time course of drug effect. Present psychopharmacologi-

cal studies are typically before and after, two-point studies, with the examiners often completely out of contact with the patient during the intervening treatment period. The only sequence handled is from before to after, as the intervening changes remain unobserved and unanalysed. As long as the field was preoccupied with the overriding question of proving therapeutic efficacy, the use of before and after controlled studies was rational and efficient. As we are now shifting to the more refined question of which drug for which patient, we must correspondingly refine and intensify our methods, possibly using modal sequential drug-change patterns. This effort will require the intensive longitudinal study of the individual patient by the trained psychiatric clinician as he remains our only presently available tool able to recognize multivariate patterned sequences, either by comparison with specified modal sequences or inductively from atypical courses. The current organization of research into studies with large numbers of patients with brief cross-sectional examinations does not encourage such work. Even the most intensive, ambitious, scientific, and well-organized programs, such as the VA and NIMH collaborative studies, do not deal directly with the question of sequential historical and drug-effect analysis, although they point in this direction.

Furthermore, it follows from the theory presented, that for many patient groups, especially those with periodic activation-affective dysregulation, the impact of drugs is to accelerate the natural course of recovery from the illness. Therefore, in the short space of several months one is given the unusual opportunity of getting an overview of a naturally much longer process, thus allowing the sharp delineation of longitudinal patterns that would otherwise be obscured by life's vicissitudes and the usual difficulties of prolonged longitudinal observation. Critical test of this hypothesis of an isomorphism between the course of both natural recovery and drug treatment

would require analyses similar to those outlined above.

These analyses would require the integrated services of a research team of skilled psychiatric clinicians, psychologists, social workers and statisticians. Furthermore, such studies could only be carried out in an inpatient setting that would welcome and support research programs, rather than consider them interferences with the clinical process. It would appear that such intensive studies will require the development of clinical research centers devoted to the specific problem of diagnosis.

IV. Summary

The practical necessity for defining the indications for psychotropic drug treatment has reemphasized that our psychiatric diagnostic system is unsatisfactory.

A series of illustrations of typical clinical responses to drugs are presented, derived from clinical and experimental experience at Hillside Hospital. These illustrations lead to the following inferences:

1. Certain psychiatric illnesses (manic-depressive and hyperactive paranoid) are primarily caused by deranged affective-activation mechanisms and have minimal secondary, functionally autonomous, symptomatic reverberations. These illnesses have naturally occurring complete remissions and also respond best to psychotropic drugs.

2. Other psychiatric illnesses (periodic phobic-anxious states, agitated depressions in obsessional states, emotionally unstable character disorders, and the schizophrenias) are characterized both by affective-activation disorders and a secondary residue that persists after the termination of the affective derangement. Psychotropic drugs again mimic the natural course of the illness since these illnesses are characterized by naturally occurring remissions with persistent secondary symptoms.

3. Other psychiatric illnesses that are not characterized by a primary activation-

affective disorder are refractory to psychotropic agents. However, histrionic-hysterical patients and chronic anticipatory anxiety states require special differentiation since their affective difficulties appear to have a different origin.

4. Psychotropic drugs may work by normalizing deranged affective-activation mechanisms. To the degree that a psychiatric illness is due to such deranged activation mechanisms, the drug effects will mimic the processes of natural remission.

5. The prediction of drug effects from target symptoms from effects on diagnosis-irrelevant control mechanisms and from specific effects on pathological states is discussed.

6. The need for study of patterned multivariate diagnosis utilizing longitudinal historical and drug effect data is emphasized. This will require much more intensive studies than were necessary for the determination of the therapeutic efficacy of psychotropic drugs. Such studies require the development of clinical research centers devoted to the specific issue of diagnosis.

REFERENCES

1. KLEIN, D. F., and FINK, M.: "Psychiatric Reaction Patterns to Imipramine." *Amer. J. Psychiat.* 119, 432-438, 1962.
2. KLEIN, D. F., and FINK, M.: "Behavioral Reaction Patterns to Phenothiazines." *Arch. Gen. Psych.* 7, 449-459, 1962.
3. KLEIN, D. F., and FINK, M.: "Multiple Item Factors as Change Measures in Psychopharmacology." *Psychopharmacologia*, 4, 43-52, 1963.
4. BELMONT, I., POLLACK, M., WILLNER, A., KLEIN, D. F. and FINK, M.: "The Effects of Imipramine and Chlorpromazine on Perceptual Analytic Ability, Perceptual Responsivity and Memory as Revealed in Rorschach Responses." *J. Nerv. & Ment. Dis.* 137, 42-50, 1963.
5. POLLACK, M., KLEIN, D. F., WILLNER, A., BLUMBERG, A. G. and FINK, M.: "Imipramine-Induced Behavioral Disorganization in Schizophrenic Patients: Physiologic and Psychologic Correlates." *Rec. Adv. Biol. Psychiat.* 7, 53-61, 1965.
6. FINK, M., KLEIN, D. F. and KRAMER, J. C.: "Clinical Efficacy of Chlorpromazine-Procyclidine Combination, Imipramine and Placebo in Depres-

sive Disorders." *Psychopharmacologia*, 7, 27-36, 1965.

7. FINK, M., POLLACK, M., KLEIN, D. F., BLUMBERG, A. G., BELMONT, I., KARP, E., KRAMER, J. C. and WILLNER, A.: "Comparative Studies of Chlorpromazine and Imipramine. I: Drug Discriminating Patterns." *Neuropsychopharmacology*, 3, 370-372, 1964.

8. POLLACK, M., KARP, E., BELMONT, I., WILLNER, A., KLEIN, D. F. and FINK, M.: "Comparative Studies of Chlorpromazine and Imipramine. II: Psychological Performance Profiles." *Neuropsychopharmacology*, 3, 374-376, 1964.

9. BLUMBERG, A. G., KLEIN, D. F., and POLLACK, M.: "Effects of Chlorpromazine and Imipramine on Systolic Blood Pressure in Psychiatric Patients: Relationships to Age, Diagnosis and Initial Blood Pressure." *J. Psychiat. Res.* 2, 51-60, 1964.

10. LEVENSTEIN, S., KLEIN, D. F., and POLLACK, M.: "Follow-up Study of Formerly Hospitalized Voluntary Psychiatric Patients: The First Two Years." (As submitted to *Amer. J. Psychiat.*)

11. LEDLEY, R. S., and LUSTED, L. B.: "Reasoning Foundations of Medical Diagnosis." *Science*, 130, 9-21, 1959.

12. KLEIN, D. F.: "Delineation of Two-Drug Responsive Anxiety Syndromes." *Psychopharmacologia*, 5, 397-408, 1964.

13. KLEIN, D. F.: "Behavioral Effects of Imipramine and Phenothiazines, Implications for a Psychiatric Pathogenetic Theory and Theory of Drug Action." *Rec. Adv. Biol. Psychiat.* 1, 273-287, 1965.

14. DUFFY, E.: *Activation and Behavior*. John Wiley & Sons, Inc. New York, 1962.

15. POLLACK, M.: "Comparison of Childhood, Adolescent and Adult Schizophrenia." *Arch. Gen. Psychiat.* 2, 652-660, 1960.

16. BRUNE, G. G., and HIMWICH, H. E.: "Biogenic Amines and Behavior in Schizophrenic Patients." *Rec. Adv. Biol. Psychiat.* 5, 144-160, 1963.

17. KELLY, D. H. W., and SARGANT, W.: "Present Treatment of Schizophrenia. A Controlled Follow-up Study." *Brit. Med. J.* 1, 147-150, 1965.

18. WARNER, H. R., TORONTO, A. F., VEASEY, L. G., and STEPHENSON, R.: "A Mathematical Approach to Medical Diagnosis." *J. Amer. Med. Assoc.* 177, 177-183, 1961.

Supplementary Comments by Dr. Klein as Presented at the Conference

I hope that presenting my paper will be redundant for you so I will not do so. I

will review the ideas of the paper in relation to this conference.

The methodological highlight of the conference, for me, was Dr. Torgerson's paper wherein he pointed out that the typological and polydimensional models were not mutually exclusive, and it was quite possible that in reality mental illness consisted of distinct types associated with common dimensions of variation.

Each of our measures then are an aspect of an underlying structure. I believe that data configurations may rule out certain underlying structures, but that an indefinitely large number of structures may be capable of producing a single data array and therefore, it is hopeless to expect any mathematical method to deduce the correct structure from our data arrays. That is, I feel that the mixture problem is in principle unsolvable without covert theory.

Postulating a specific structure is simply hypergeometric phraseology for forming a theory; that is, the specification of relevant constructs and the stipulation of the propositional relationships of these constructs.

Now my paper proposes that there are a variety of qualitatively unique psychiatric illnesses probably derived from different etiologies. However, I believe that some of these types share a dimension, clinically well known and frequently discovered in factorial or typological studies—the bipolar dimension of retardation and withdrawal versus mania and excitation. Patients with qualitatively distinct disorders may nevertheless be extremely difficult to distinguish on the basis of cross-sectional behavior. If they are in a state of extreme activation-dysregulation it is very hard clinically, or by any clustering method, to tell an excited schizophrenic from an acute manic, or a retarded withdrawn perplexed schizophrenic from a severe retarded stuporous depression, on the basis of interview data.

If we look at the typologies presented to us by Lorr, Grinker, Katz, Overall, and Gerard, we find in all cases the dimen-

sion of retardation versus excitation prominent in the definition of the categories. Dr. Katz reminded us of the introversion-extroversion dimension as perhaps basic to his unexplained typological symmetry. I agree to the need for finding an explanation for typological symmetry in the form of common, bipolar dimensions, since types are nominal constructs and symmetry indicates underlying relationships.

Furthermore, if we look at Dr. Lorr's types, since he was kind enough to compare his typology with classical nosology, the excited type cuts across both manics and paranoid schizophrenics. As a clinician who is used to longitudinal assay, I cannot help but feel that his excited, excited-hostile, and hostile-paranoid groups merely represent different stages of the same disorder, fluctuating in degrees of activation. Certainly the demographic qualities of these three groups, as presented by him, are not strikingly different.

You may wonder what happened to my topic, psychotropic drugs. I believe that it is precisely the physiological derangement of the central regulating agency that governs the individual's position on this dimension that is responsive to the phenothiazines and antidepressants and that, in fact, may be the only effect of these drugs.

If this is so, why haven't the attempts to predict drug effect from an initial behavioral status assaying this dimension been more convincing? I would suggest several reasons.

First, it is easier to get a discriminant function, canonical correlation, or multiple regression equation than it is to get a scattergram. It is my contention that the relationship of behavioral indices to outcome is highly nonlinear and that these effects have not been allowed for, in most analyses. For instance, if we take the standard bipolar activation dimension and compare it with the global measure of improvement on chlorpromazine, we get a U-shaped bivariate distribution. That is,

both the retarded and the excited do well; the intermediate group do poorly.

Second, I believe that apparently monopolar scales are what I have facetiously called cryptobipolar or perhaps pseudomonopolar. Now what do I mean by that?

A scale such as "hostility" may go from "no hostility" through "abusive" to "assaultive." The assaultive end refers to a high portion of overactivated people, but the "no hostility" end mixes those who are normally activated and not hostile with those who are not hostile because of other pathological phenomena. Retardation is preventing hostility.

When such pseudomonopolar scales are scattered against outcome measures we regularly see a triangular array. That is, high scores predict a good outcome, but low scores may end up either good or bad. I think that such relationships, U-shaped, triangular, et cetera, require rather special care and feeding, and should not be indiscriminately plugged into the standard multivariate models.

Finally, and this is, I believe, crucial to the entire issue of prediction of outcome, after the activation pathology is cared for by the drugs, we are left with a spread of residual pathological traits that are dependent on the basic pathological type and that these residual pathological traits cannot be predicted simply from the initial behavioral data.

They can be predicted from diagnosis which, as Dr. Klerman reminded us, is not simply the covariation of symptomatology. In my paper, which I will not review, I emphasize the difference between drug-responsive pathology and residual functionally autonomous pathology, based both on diagnostic type and new maladaptive mis-carried repairs.

Finally, I would like to repeat Dr. Brill's remark to me, that we will never talk our way out of our dilemmas, we must work our way out. I believe that this cannot be done by good will alone and it cannot be done by *ad hoc* personal relationships of

psychologists, psychiatrists, statisticians, et cetera. It requires an organizational framework. The only way you get something this complex done is by putting a team to work on it, in the proper setting.

The development of federally supported, well-staffed, clinical research centers of approximately 150 to 200 psychiatric beds would further this goal and enable us to centralize our efforts.

Patterns of Reaction to Drug Treatments Derived Through Multivariate Procedures

Dean J. Clyde, Ph. D.¹

In our laboratory we have been experimenting with procedures for classifying patients according to their response to drugs. If we are successful the results could have both practical and theoretical significance. Practical significance, because if we can predict that a patient will improve more on one drug than another, the procedure is useful clinically for selecting the right drug for the right patient. Theoretical significance, because patients who respond to a drug might be etiologically different from patients who do not respond.

We began our statistical experiments by trying better to understand the concept of drug response. What do doctors and nurses mean when they say a psychiatric patient has improved?

The data we used were gathered in eight hospitals collaborating in a study sponsored by the National Institute of Mental Health, comparing acetophenazine, chlorpromazine, and fluphenazine in the treatment of schizophrenics.

After 5 weeks of drug treatment, each patient was rated by doctors and nurses with the following question:

¹ Director, Biometric Laboratory, University of Miami, Fla.

Compared to his condition at admission to the project, how much has he changed?

- Very much improved
- ☐ 1
- Much improved
- ☐ 2
- Minimally improved
- ☐ 3
- No change
- ☐ 4
- Minimally worse
- ☐ 5
- Much worse
- ☐ 6
- Very much worse
- ☐ 7

In addition, each patient was rated before and after 5 weeks of treatment with some items from the Minnesota-Hartford Personality Assay. Doctors used 127 items and nurses used 142 items. Scoring systems were devised for both the doctor's and nurse's scales by factor analyzing the items. We selected 11 of the most conspicuous factors from the doctor's scale and 7 from the nurse's scale to see how they related to improvement.

Presumably in rating a patient's improvement on the seven-point scale, a doctor or nurse recalled how the patient was before treatment, considered his condition after 5 weeks of treatment, and evaluated the difference between the two occasions. Change in certain characteristics was undoubtedly weighted more heavily than change in other characteristics, with regard to their relevance to improvement. We wished to give quantitative expression to these weightings, and also determine whether doctors and nurses emphasize the same characteristics. In other words, what specifically do doctors and nurses mean by improvement, and do doctors and nurses mean the same thing?

Table 1 shows the mean scores given 278 patients before and after 5 weeks of treatment. On the average the doctors and

TABLE 1.—Symptoms Before and After 5 Weeks of Treatment Related to Degree of Improvement (N=278)

Time of rating	Rater	Variable	Mean	Standard deviation
Before.	Doctor.	Thinking disorder	4.6	1.6
		Happiness.....	1.9	1.0
		Hostility.....	2.8	1.4
		Conformity.....	2.6	1.0
		Guilt.....	3.0	1.7
		Compulsivity....	2.4	1.5
		Dissociation....	2.7	1.8
		Paranoid.....	4.5	2.0
		Anxiety.....	3.0	1.7
		Somatic concern .	2.0	1.5
		Fluency.....	3.1	1.6
	Nurse..	Thinking disorder	5.0	1.3
		Happiness.....	2.0	1.0
		Hostility.....	2.8	1.3
		Conformity.....	2.3	1.0
		Guilt.....	2.8	1.4
		Compulsivity....	2.8	1.2
		Tranquillity....	2.7	1.4
After..	Doctor	Thinking disorder	3.2	1.6
		Happiness.....	2.3	1.1
		Hostility.....	2.2	1.0
		Conformity.....	3.5	1.2
		Guilt.....	2.6	1.3
		Compulsivity....	2.5	1.2
		Dissociation....	1.9	1.2
		Paranoid.....	2.9	1.8
		Anxiety.....	2.3	1.4
		Somatic concern..	1.9	1.3
		Fluency.....	3.6	1.3
	Nurse..	Thinking disorder	3.6	1.5
		Happiness.....	2.5	1.0
		Hostility.....	2.4	1.2
		Conformity.....	3.0	1.1
		Guilt.....	2.8	1.2
		Compulsivity....	3.1	1.1
		Tranquillity....	3.5	1.5

Time of rating	Rater	Variable	Mean	Standard deviation
After.	Doctor.	Improvement....	2.7	1.1
	Nurse..	Improvement....	2.7	1.1

Latent root No.	Canonical correlation	P less than
1	0.84	0.0001
2	.56	.0001

nurses observed about the same amount of improvement, both yielding scores of 2.7, or somewhere between “minimally improved” and “much improved” on the seven-point scale.

Table 1 also shows the canonical correlations between the two sets of ratings, together with levels of significance. Both of the canonical correlations are very significantly different from zero. This means that before and after ratings on the Minnesota-Hartford Personality Assay are related to improvement in two distinctly different ways. There are at least two different patterns of improvement following drug treatment.

Examination of tables 2 and 3 reveals more about these patterns of improvement. In table 2 the weights for both the doctors’ and nurses’ improvement scores are posi-

tive, meaning that for some of the patients they agree as to who improved and who did not. The doctors gave the most weight to changes in thinking disorder and conformity, and the nurses gave the most weight to changes in thinking disorder, happiness, and hostility.

In table 3 the weights for the doctors’ and nurses’ improvement scores are opposite in sign, meaning that for some of the patients they did not agree as to who improved. For this group of patients the doctors gave the most weight to changes in thinking disorder, conformity, and somatic concern, and the nurses gave the most weight to changes in conformity and tranquillity.

In order to clarify the two different patterns of improvement, we selected the five patients who showed the most improvement

TABLE 2.—Weights Associated With Latent Root No. 1

Time of rating	Rater	Variable	Weight
Before	Doctor.	Thinking disorder .	−0.012
		Happiness180
		Hostility	−.007
		Conformity158
		Guilt026
		Compulsivity130
		Dissociation	−.145
		Paranoid007
		Anxiety114
		Somatic concern . .	−.046
	Nurse..	Fluency	−.020
		Thinking disorder .	−.148
		Happiness134
		Hostility	−.241
		Conformity035
		Guilt034
		Compulsivity	−.047
		Tranquillity	−.041
After	Doctor.	Thinking disorder .	.260
		Happiness	−.075
		Hostility	−.072
		Conformity	−.292
		Guilt017
		Compulsivity	−.009
		Dissociation	−.057
		Paranoid035
		Anxiety019
		Somatic concern . .	.009
	Nurse..	Fluency018
		Thinking disorder .	.302
		Happiness	−.226
		Hostility353
		Conformity	−.137
		Guilt	−.044
		Compulsivity	−.022
		Tranquillity	−.066

Time of rating	Rater	Variable	Weight
After	Doctor.	Improvement471
	Nurse..	Improvement667

TABLE 3.—Weights Associated With Latent Root No. 2

Time of rating	Rater	Variable	Weight	Time of rating	Rater	Variable	Weight
Before	Doctor.	Thinking disorder .	0.074	After	Doctor.	Improvement	−1.082
		Happiness	−.156			Improvement974
		Hostility	−.282				
		Conformity	−.345				
		Guilt107				
		Compulsivity	−.034				
		Dissociation	−.112				
		Paranoid125				
		Anxiety151				
		Somatic concern083				
	Nurse..	Fluency113				
		Thinking disorder .	−.140				
		Happiness190				
		Hostility007				
		Conformity017				
		Guilt063				
		Compulsivity254				
		Tranquillity	−.333				
		Thinking disorder .	−.310				
		Happiness156				
After	Doctor.	Hostility	−.202				
		Conformity404				
		Guilt257				
		Compulsivity027				
		Dissociation	−.138				
		Paranoid073				
		Anxiety	−.219				
		Somatic concern . . .	−.295				
		Fluency112				
		Thinking disorder .	.280				
	Nurse..	Happiness	−.172				
		Hostility138				
		Conformity	−.383				
		Guilt201				
		Compulsivity	−.213				
		Tranquillity	−.059				

of each type. Figure 1 displays mean scores before and after 5 weeks of treatment for five patients rated “very much improved” by both doctor and nurse.

Figure 2 shows mean scores before and after treatment for five patients rated “very much improved” by doctors but “minimally improved” by nurses. On none of the 278 patients was there a disagreement greater than this.

I draw the following conclusions from tables 1, 2, and 3 and figures 1 and 2: Detailed ratings using items from the Minnesota-Hartford Personality Assay reveal two patterns of change following drug treatment. These two patterns can be thought of as two types of patients, two subgroups within the total group labeled “schizophrenia.” There might be more

than two subgroups; the mathematical technique does not answer that question. Doctors and nurses agree that one type of change is improvement. They also agree that the other type of change occurs, but disagree as to whether it should be called improvement or not.

The two patterns of change do not seem to me to correspond to any subtypes of schizophrenics identified by previous diagnostic systems. Drug response appears to cut across traditional diagnostic groupings and yield a new basis for classifying psychiatric patients. I do not know exactly what to make of this.

The next question we asked was: Can we predict which patients will improve on drugs with respect to either of these two patterns of change? Given a patient’s

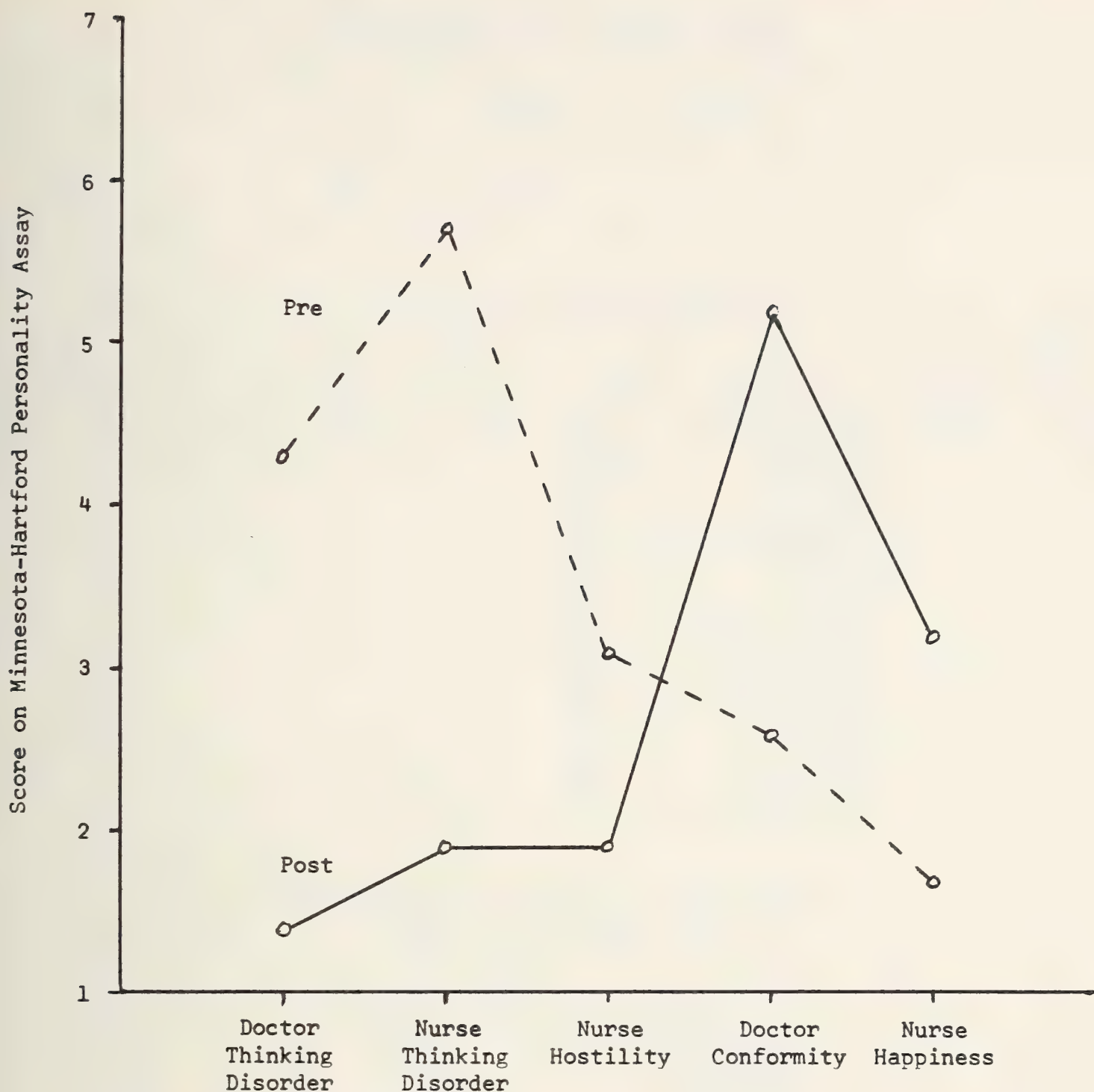


FIGURE 1.—Mean scores of five patients rated "very much improved" by both doctors and nurses.

scores on the Minnesota-Hartford Personality Assay before treatment, what can we say about his future drug response?

Table 4 shows the canonical correlations between pretreatment ratings and improvement after 5 weeks of treatment. We can predict better than chance for only one type of improvement—the type on which doctors and nurses disagree. The correlation of 0.35 is significant at the 0.04 level.

Evidence from other studies indicates that additional variables would increase the accuracy of our prediction. It is interesting that symptom ratings alone are apparently of some value.

Three drugs were employed in the study from which these data were taken. Does one drug produce more of this peculiar type of predictable improvement than another?

Table 5 gives the analysis of variance of composite drug response scores, weighted according to the correlation from table 4. This pattern of change is called improvement more by the doctors than the nurses, and is predictable to some extent from pre-treatment symptoms. The drugs differ very significantly with p less than 0.0001. Acetophenazine produced much more of

TABLE 4.—Symptoms Before Treatment Related to Degree of Improvement After Treatment (N=278)

Latent root No.	Canonical correlation	<i>P</i> less than
1	0.35	0.04
2	.26	.38

Weights Associated With Latent Root No. 1

Time of rating	Rater	Variable	Weight	Time of rating	Rater	Variable	Weight
Before	Doctor.	Thinking disorder .	−0.178	After	Doctor.	Improvement	−1.179
		Happiness	−.186			Improvement581
		Hostility	−.629				
		Conformity	−.042				
		Guilt135				
		Compulsivity	−.174				
		Dissociation	−.079				
		Paranoid100				
		Anxiety021				
		Somatic concern179				
	Nurse..	Fluency	−.020				
		Thinking disorder .	.076				
		Happiness231				
		Hostility	−.072				
		Conformity	−.043				
		Guilt317				
		Compulsivity381				
		Tranquillity	−.833				

TABLE 5.—Doctors' and Nurses' Composite Improvement Score, Using Weights From Table 4

Treatment	Number of patients	Mean score
Acetophenazine	96	0.39
Chlorpromazine	84	−.24
Fluphenazine	98	−.18

Analysis of Variance

Source of variation	Degrees of freedom	Sum of squares	Mean square	<i>F</i>	<i>P</i> less than
Between drugs	2	23.01	11.51	12.51	0.0001
Within	275	253.94	.92		
Total	277	276.95			

this type of response than either chlorpromazine or fluphenazine.

When we examined the patients on whom the doctors and nurses agreed as to improvement, we found that acetophenazine appeared only slightly better than the other two drugs, with *p* less than 0.04.

To summarize our findings, we have discovered a common thread running through all of our analyses. When we studied what doctors and nurses mean by improvement, we identified a minority group of patients for whom doctors and nurses agreed as to change but disagreed as to improvement.

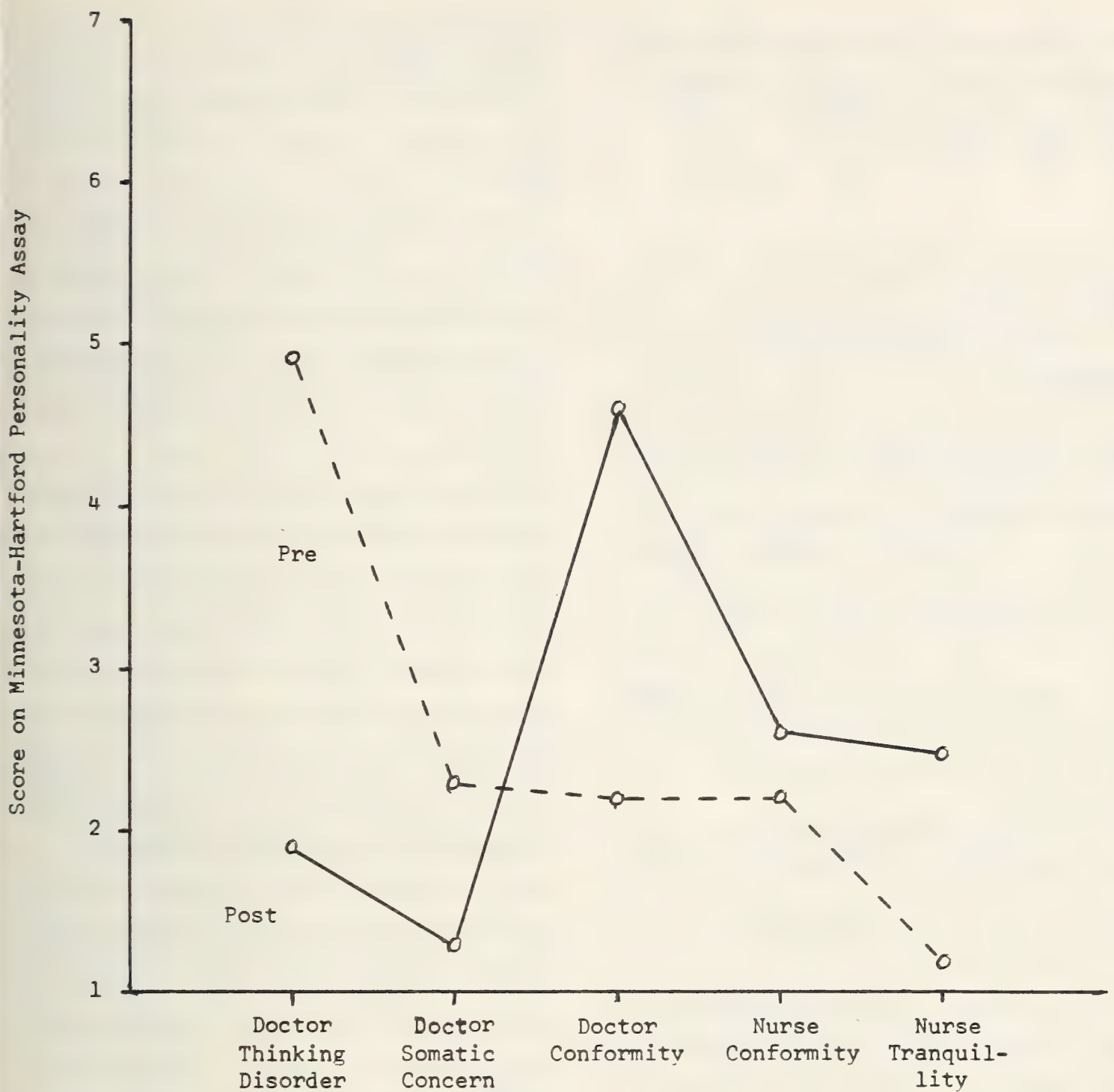


FIGURE 2.—Mean scores of five patients rated "very much improved" by doctors but "minimally improved" by nurses.

When we looked at predictors, we found that the drug response of this minority group was predictable to some extent from pretreatment symptom ratings. When we compared the three drugs, we found that this same group benefited quite a bit more from acetophenazine than from the other two drugs.

There is no reason I can think of except one for the appearance and reappearance of this minority group of patients in entirely separate and independent types of analyses: the group really exists.

Perhaps for too long we have been focussing on patients who everyone agrees have benefited from drugs. There appears to be another group on whom professional observers regularly disagree, and acetophenazine is most efficacious for them.

The analyses presented here give us a rather strange and unsatisfactory basis for diagnosing this type of schizophrenia: doctors and nurses disagree as to how much the patients benefit from drugs. Such a diagnosis can be made only after a period of drug treatment. There must be better

ways of diagnosing this group prior to drug treatment and perhaps gaining insight into etiology. We are continuing our explorations along these lines.

Supplementary Comments by Dr. Clyde as Presented at the Conference

One of the purposes of this conference might be stated thus:

Can statisticians give any advice to clinicians on solving the problem of classification in psychiatry? I am a middleman between statisticians and clinicians, and I will answer "Yes" to that question. There are a number of points at which statisticians can give good advice to clinicians on this problem. I will discuss just one point this morning.

I'll try to get this one point across to you rather than make it as an assertion. I'll try to show you the reasoning behind this bit of advice that I think statisticians would give clinicians.

I have drawn a scatter diagram on the board. This is to show the correlation between height and weight. Each dot on the board represents one of the people attending this conference.

[Laughter.]

I am picking a very simple and obvious example to try to illustrate my analogy. Suppose we wish to find out if there is a relationship between height and weight by scientific methods, and suppose we are members of a committee to plan the study. In this committee meeting a number of people speak up and say, "Well, we just can't be sloppy and pick any people to determine this relationship. We have to be very scientific. We have to pick homogeneous groups. We have to narrow down the focus on the specific problem and study this very intensively in order to get the answer. We will pick only people between

71 and 72 inches tall and further, to be very scientific and precise, we will pick people between 155 and 160 pounds." So we are going to study the people in this cell.

First of all, of course, we drastically reduced our number of cases. In addition, the distribution of points within this cell is almost random even though, if we took the whole group, we'd see that the distribution of points in the scatter diagram is not random.

But if we focus down on this very homogeneous group and are terribly scientific and compute the correlation coefficient between height and weight in this cell, we might get something like 0.06, which is not statistically significant for the number of cases involved. So we will solemnly announce that after this very careful scientific study in which we picked homogeneous groups and carefully defined our sample, we have found that there is no significant relationship between height and weight.

To be sure of our conclusion we will replicate the study. We will go to another conference that is being held in this building and measure the peoples' heights and weights, and we will pick a very homogeneous subgroup and come to the same conclusion.

One of the points I am making is that it is possible to follow a fallacious procedure over and over and arrive at the same fallacious conclusion. The mere fact that we get consistent results over time only proves that it is possible to repeat our mistakes.

If we wish to study the relationship between two variables we should not unduly restrict the range of them by selecting only homogeneous subgroups to study.

Let me erase these headings "Height" and "Weight" on the board and instead write in "Psychopathology" and "Drugs."

We wish to know if there is a relationship between psychopathology and drug effect. We want to know if certain of the drugs available to us have certain effects

on psychopathology, or conversely, we want to know if patients with certain psychopathological symptoms react differently to different drugs.

In our committee meeting, therefore, we decide that we must not include a wide range of psychopathology and drugs; we must select a very homogeneous group of patients, excluding all sorts of people for various reasons in order to focus on a specific problem and then we must also select rather similar drugs to be studied.

It's possible that with very elaborate statistical techniques we can show within this narrow range that there is a relation between psychopathology and drug effects. But we're certainly stacking the cards against ourselves and making it very difficult to detect a relationship if it exists.

I'm happy to say that I think I see a trend to include a broader range of psychopathology in our studies and also a broader range of drugs. There may be very clear underlying relationships between the types of drugs we have available and the types of psychopathology that are affected by them. If we study broader ranges of these things we are much more likely to get positive results.

In summary I think that one piece of advice a statistician would give to clinicians is that if you want to study the relationship between two things, do not unduly restrict the range of either of them. If you do you might fail to find a relationship which should be perfectly evident.

I would like to suggest that the next time you are in a committee meeting planning a study and you find yourself on the verge of suggesting that we must focus on the problem very specifically, that we must select very carefully a precisely defined homogeneous group for our study—before you say that stop and think: Would a professional statistician agree with this idea?

You may be on the verge of speaking nonsense.

Discussion

ARDIE LUBIN, Ph. D., *Research Psychologist, U.S. Navy Medical Neuropsychiatric Research Unit, San Diego, Chairman*

(Papers: "Process and Reactive Schizophrenia—Some Conceptions and Issues," Norman Garmezy, Ph. D., Director, Center for Personality Research, Department of Psychology, University of Minnesota. "Diagnosis and Pattern of Reaction to Drug Treatment—Clinically Derived Formulations," Donald F. Klein, M.D., Director of Research, Hillside Hospital, Glen Oaks, N.Y. "Patterns of Reaction to Drug Treatments Derived Through Multivariate Procedures," Dean J. Clyde, Ph. D., Director, Biometric Laboratory, University of Miami.)

FORMAL DISCUSSION OF PAPERS

	Page
HEINZ E. LEHMANN, M.D.	491
C. JAMES KLETT, Ph. D.	493

OPEN DISCUSSION

Drs. GLUECK, KLEIN, CLYDE, KLETT, GARMEZY, GRUENBERG, COLE, KLERMAN, and ZUBIN.....	500
---	-----

Discussant's Remarks

HEINZ E. LEHMANN, M.D., *Clinical Director, Douglas Hospital, Verdun; Professor of Psychiatry, McGill University, Montreal, Canada.*

Dr. Garmezy's paper will have to be read. It is a fine synthesis of reference material which everyone working in the field of schizophrenia should read.

Dr. Garmezy has focused on the old and by no means exhausted controversies of reactive versus process, endogenous versus exogenous, organic versus psychogenic, good prognosis versus bad prognosis. They are not the same but they are all somehow related, and they are, of course, the very lifeblood of the clinician.

Now, endogenous-exogenous: there is a great deal of philosophical reluctance to accept this distinction because how do we know that the causes for a disorder are

endogenous? Perhaps they are exogenous and we simply can't see the causative factors.

From the clinician's point of view I still think this is an excellent distinction. If we rely on an operational definition that reactive is that for which we can readily determine a cause or a traumatic incident that would lead to the particular reaction under observation—if we accept this operational definition, then we can use this term reactive predictively as a powerful concept by saying for instance that a good prognosis is usually associated with those reactive cases where we can readily identify a cause. But in those cases where in a schizophrenic patient we cannot readily identify a traumatic incident or event, physical or psychogenic, we call it endogenous by operational definition and we do know that the outcome in most of these cases is much poorer, or at least the risk is greater for a poor outcome.

What about organic versus psychogenic? I think I had better keep out of this. I don't know whether this controversy will ever come to some sort of satisfactory solution because quite likely there is not really a clear distinction between organic and psychic. Again we come to the same impasse, how much can we readily identify as organic and how much can we not? And accordingly we will have to make our operational definition.

Now the diagnosis—to come back briefly to what Dr. Klerman mentioned yesterday—is of course what the clinician is not only interested in but deeply concerned with. He has to know not only what a particular pathological configuration which is present at the moment reflects, but also what is likely to happen later and what he is going to do about it. And what he is going to do about it, of course, depends on the response tendencies of his patient. The response tendencies, in turn, depend on the history of the patient, the immediate history or the immediate traumatic event and also on the long-ago history. What kind of

a family setup has there been? What kind of a family setup is there now?

Allow me to spend a minute on what is actually going on in a diagnostic conference: The patient is presented and the doctor who presents the case recites the symptoms, describes the patient more or less accurately, and the diagnosis is not difficult to make in most cases. This may be evidently a paranoid schizophrenic. He expresses delusions of persecution, or is suspicious and withdrawn and somewhat peculiar and certainly is not socializing well, and so on. He does not sleep well, and the nurses might report a few other things that they have observed.

So the diagnosis is made. Now the next question is, how long has this been present? If it has been present a month or two, I might immediately say: "In another month or two this patient should be well and leave the hospital, and if he doesn't then there is something wrong with our treatment. We ought to be able to get him (or her) well in 2 months."

If, on the other hand, the patient has been paranoid for the last 2 or 3 years, I would say: "Well, it will probably take 5 to 6 months." This or something similar is automatically done in most clinical setups, I think. Yet it is not included in most of the rating scales—the duration of the illness prior to examination.

Then we would ask: What kind of a family has he had? Perhaps it was an awful family setup. The mother was never home and the father was usually drunk, and the parents have been separated for a long time, and the child was unwanted, and—well, that makes the prognosis quite a bit poorer now. Our prognostic rating goes down.

And what kind of a family setup has the patient now? Well, there is a husband and he is quite well motivated and he wants her back and he and the patient always had a good relationship. Well, that makes the prognosis a little better. But it may not

quite neutralize the bad effects of the childhood setup.

And then somebody will say, "Well, I saw her the first day she was brought in—she wouldn't talk and she was obviously very suspicious of everyone. And when I told her that she is all mixed up and that she can't think straight, that she is mentally ill and that she is here to be helped, and that we are quite sure we can help her, she seemed to look as though she was almost happy about it, and it was quite evident that she felt relieved by being told that she was mentally ill and that she could be helped."

That is again a point in favor and our prognostic rating goes up.

Taking all these factors into consideration, we now have determined that probably instead of 1 or 2 months it will take 3 or 4 months but by that time the patient should be well, and I would think in about 80 percent of the cases, such predictions after such conferences will be justified and will turn out to be fairly valid.

However, you see, there are a lot of things which go into this which are not in the rating scales and not in interview behavior.

I see that I have 2 minutes left. May I now just spend 1 minute on Dr. Klein's paper which to me appears as a stimulating paper for any clinician because while he is recommending clustering he thinks of clustering according to systems which come from outside the symptoms. In other words, he has an activating system, an effective system, a regulatory system. And if you want to, you may put these systems into the limbic lobe and into the cortex and into the reticular activating system. But these are cardinal systems, one might say, and the clusters of symptoms are ordered around the systems rather than just around their own statistical distribution.

And Dr. Klein proposes that drugs or any kind of treatment should be aimed at treating systems or clusters of symptoms according to systems.

Dr. Clyde's paper is quite a teaser for a clinician and also, I think, for the experimentalists and statisticians because it indicates very clearly that what is good for the goose is not good for the gander. What the doctor would consider improvement the nurse does not consider improvement, and what the doctor might see as a most important symptom prior to treatment the nurse might not see as so important.

This does not simply mean that they have different approaches to their patients. It also means that one can base actual predictions on it. If, for instance, the nurse does not see much hostility in her patients, then these patients will probably not do as well with acetophenazine as they would do if the nurse would have seen more hostility.

And if the nurse does not see a thinking disorder in the patient before treatment and only the doctor sees it, again the prognosis is not as good. Whatever this means, we will simply have to consider carefully who is doing the rating—whether the nurse, experienced or inexperienced, a psychiatrist, or a psychologist—and according to who rates and what they see, the prognosis might well be different.

Discussant's Remarks

C. JAMES KLETT, Ph. D., *Chief, Central Neuropsychiatric Research Laboratory, Veterans Administration Hospital, Perry Point, Md.*

I have to respond to these three papers from what I recognize to be a rather narrow point of view; namely, classification as a stratagem of research in psychopharmacology. In doing so I am considering classification in a much wider sense than psychiatric diagnosis. Obviously formal psychiatric diagnosis serves many purposes, some of them more effectively than others. In my experience it has not been particularly helpful in evaluating the

clinical effectiveness of drugs. Because so many other individuals presenting alternative classifications at this conference are involved in one way or another in drug evaluations, I would judge that they too have found the current diagnostic system to be of limited value in this context.

The need for a classification of patients in a therapeutic trial arises in a number of ways. All patients do not respond to treatment. Can we identify those that will and those that won't? It has been observed that some patients respond to one drug but not another. Can it be that drugs have a selective effect, working with a certain patient class while having no effect on another class? Is there a drug of choice? The right drug for the right patient? Some patients relapse quickly when drugs are withdrawn. Can they be described as a type that differs from another type that relapses slowly or not at all? Speed of response to drugs or the appearance of certain side effects could also lead to a search for types. In all of these questions which are raised by the variability of patients on the various parameters of drug response, the goal is to predict outcome of treatment. This needn't involve the concept of type at all, although this has often been the direction research has taken.

In the usual study of the relative effectiveness of two drugs, patients are randomly assigned to treatment and observations are recorded before and after a fixed interval of drug administration. Usually there is considerable variation in response and usually when working with clinically effective compounds, no significant difference is found between groups. Subclassification of the patient sample is often introduced in this design for two reasons: To sharpen up the test of differences between drugs and to test for a drug-by-patient-class interaction. Obviously though some way of subclassifying patients has to be found that is related to drug response. The three papers by Garmezy, Klein, and Clyde illustrate, as do many other papers of this con-

ference, the diversity of the attempts that have been made to find satisfactory classification systems.

As Dr. Garmezy's review illustrates, the concept of process and reactive schizophrenia has been a very persistent one in psychiatric thought. Whether this concept is used as a dichotomy or a continuum, it does provide a focus around which can be organized a great many variables which have been found to show a low-order relationship with outcome. Each of the favorable and unfavorable prognostic indices included in his tables 1 and 2 or similar lists constructed by other investigators (1, 2, 3) could be used individually as a way of subclassifying patients in a therapeutic trial. Many of them have been used in this way. However, none of these variables is highly correlated with outcome, most of them are intercorrelated, and there are so many of them that it is not satisfactory from a statistical point of view to deal with them individually as independent variables. The Phillips Premorbid Scale and the Elgin Prognostic Scale combine some of these variables into a single score, some degree of objectivity, and quantification is gained, and the scales with the appropriate cutting scores provide an operational definition of the patient types that can be used in experiments and replicated by others. However, as Dr. Garmezy himself points out, neither of these two scales is a completely satisfactory instrument. It surprises me that a more sophisticated and comprehensive prognostic scale has not been produced.

Dr. Klein's paper reflects careful observation and clinical insight as well as an appreciation for the methodological problems involved in trying to identify and establish differential drug response in a diverse patient group. The strength of his approach to the formulation of types is that they are derived from direct study of patients' response to drugs by skilled clinicians guided by a theory of drug action rather than by purely empirical statistical

manipulation of rating scale data. However, from the standpoint of general application to research at this time, this is also the weakness of his approach. Although most clinicians might recognize his patient types, they are not sufficiently well defined in the objective and quantitative sense that they could be used successfully by very many investigators. From the same pragmatic point of view, it is difficult to see how Dr. Clyde's approach will have much utility in answering any of the important questions about patients' response to drugs.

In our own work, we have tried many approaches and a variety of statistical techniques. The Phillips Premorbid Scale and the Elgin Prognostic Scale were used in one study without success (4). We have tried a range of prognostic variables individually or in statistical combinations to try to answer several questions. We were moderately successful in predicting outcome by multiple-regression analysis over a three-year period (5) but unable to discriminate between relapsers and nonrelapsers using discriminate analysis (6). We are currently using the types described by Dr. Lorr earlier to test the difference between chlorpromazine, fluphenazine, and thioridazine within each type.

We have also tackled the problem of finding the right drug for the right patient (7). Data on the response of recently hospitalized schizophrenic patients to chlorpromazine, fluphenazine, and thioridazine were available from two large but independent multihospital studies. The Veterans Administration's cooperative study evaluated these compounds over an 8-week period (4). The National Institute of Mental Health, Psychopharmacology Service Center (NIMH-PSC) in the first of their collaborative studies evaluated them over a 6-week period (8). Although there are many differences between these two studies, they are in most respects roughly comparable. Both studies used the Inpatient Multidimensional Psychiatric Scale

(IMPS) which yields 10 syndrome scores and 3 higher-order morbidity scores which are a linear compound of them. The latter have been named in terms of the syndromes with which they are most highly correlated: excitement versus retardation, schizophrenic disorganization, and paranoid process.

Using the 74 patients who received chlorpromazine in the VA study, multiple-regression equations were developed to predict drug response from the 10 pretreatment syndrome scores. Separate equations were developed to estimate each of three posttreatment morbidity scores used as a criteria of drug effectiveness. In the same manner, 3 equations were developed for the 71 patients who had received fluphenazine, and 3 for the 76 patients who had received thioridazine. The standard regression coefficients and multiple-correlations associated with these equations are shown in table 1.

Since the pretreatment scores of any newly admitted patient could be used with these equations to estimate how well he would respond to subsequent treatment, it was reasoned that the drug associated with the equation yielding the most favorable prediction could be considered the drug of choice. Validation of this view was obtained by applying the equations developed on the VA sample to the 320 patients who had received drugs in the NIMH-PSC study. Predictions of posttreatment status on each of the three outcome criteria were made for all three sets of drug equations. For each criterion, each patient had three predicted scores which were interpreted as indicating how well he would have responded if he had received chlorpromazine, fluphenazine, or thioridazine. Considering the drugs in pairs, each patient was then assigned to his drug of choice. Since in the NIMH-PSC study, patients were assigned to these drugs at random, some of them did and some did not actually receive the drug of choice.

TABLE 1.—Standard Regression Coefficients and Multiple Correlations Developed From the VA Sample
[Decimal omitted]

Criterion—postfactor score	Drug	Predictors—pretreatment syndrome ¹										Multiple correlation
		EXC	PAR	DIS	INP	PCP	MTR	HOS	RTD	GRN	CNP	
Excitement versus retardation.	Chlorpromazine.....	49	10	27	09	-03	13	24	09	-08	-06	66
	Fluphenazine.....	07	47	-01	-01	-30	13	10	-26	-27	26	67
	Thioridazine.....	32	-01	00	-14	01	12	06	-23	22	12	63
Schizophrenic disorganization.	Chlorpromazine.....	46	04	22	19	-28	31	-16	27	-04	09	60
	Fluphenazine.....	-03	06	-13	-07	-07	37	-26	24	-22	19	65
	Thioridazine.....	-01	08	-30	-22	12	-02	-08	15	-07	26	63
Paranoid process.....	Chlorpromazine.....	-15	14	08	17	29	23	06	-08	13	13	66
	Fluphenazine.....	-42	05	00	19	-10	29	46	-21	12	-14	50
	Thioridazine.....	-41	16	00	05	39	20	14	-11	27	-21	66

¹ EXC==excitement, PAR==paranoid projection, DIS==disorientation, INP== anxious introjectiveness, PCP==perceptual distortion, MTR==motor disturbances, HOS==hostile belligerence, RTD==retardation and apathy, GRN==grandiose expansiveness, CNP==conceptual disorganization.

TABLE 2.—Actual Outcomes for the Chlorpromazine-Fluphenazine Groups in the NIMH—PSC Sample 1

Criterion	Adjusted postmeans ²			Analysis of covariance significance test	
	Actually Received	Predicted		Effect	P
		Chlorpromazine	Fluphenazine		
Excitement versus retardation	Chlorpromazine. . . .	(65) 27.2	(33) 41.6	Actual.	n.s.
	Fluphenazine.	(65) 40.1	(33) 10.0	predicted. interaction.	n.s. 0.025
Schizophrenic disorganization	Chlorpromazine. . . .	(56) 65.3	(56) 68.2	Actual.	n.s.
	Fluphenazine.	(49) 82.3	(49) 51.0	predicted. interaction.	n.s. 0.01
Paranoid process	Chlorpromazine. . . .	(62) 130.0	(41) 172.3	Actual.	n.s.
	Fluphenazine.	(62) 142.5	(41) 136.9	predicted. interaction.	n.s. 0.01

¹ Values in parentheses indicate the number of cases in each cell.

² The lower the score the greater the improvement.

For example, considering just the chlorpromazine-fluphenazine comparison, some patients who should have received chlorpromazine actually did but others received fluphenazine because of random assignment. Similarly, some patients who should have received fluphenazine actually did but others received chlorpromazine because of random assignment. The expectation was that those patients who actually received their drug of choice would show a response superior to those who were assigned at random to the other drug. The test of this hypothesis could be made in terms of the actual outcomes of the NIMH-PSC study.

Table 2 shows the average chlorpromazine-fluphenazine criterion scores adjusted for differences in initial status by covariance analysis. The hypothesis of differential drug action was confirmed by the significant interactions. Patients receiving their drug of choice showed superior response to treatment on all three criteria. No significant difference in improvement for any of the three criteria was observed for the main effects. These findings emphasize the importance of considering selective action in drug evaluation, for simple random assignment of patients would have led to the conclusion of the parent studies that chlorpromazine and fluphenazine are not significantly different in their action.

Having developed equations that predict drug action and having shown that subgroups formed by application of these equations are significantly different in the predicted direction for at least two of three drugs, it would be useful to be able to characterize the responders and nonresponders in each drug group, as well as to identify the differences between patients assigned to different drug groups. It is at this point of interpretation that we begin thinking in terms of patient types although it is not necessary to do so. In the reduction of schizophrenic disorganization the responders to chlorpromazine could be characterized by low scores particularly on EXC, DIS, MTR, and RTD, and high

scores on PCP. Responders to fluphenazine would be expected to have low scores on MTR and RTD, and high scores on HOS and GRN. However, assignment of patients to one of these drugs as the drug of choice would be most strongly influenced by the variables with greatest difference in the coefficients. Fluphenazine is favored for patients having high scores on EXC, DIS, and INP and low scores on PCP; the other syndrome scores although useful in predicting response to a drug do not contribute as much to a differential prediction.

REFERENCES

- (1) PASCAL, G. R., SWENSEN, C. H., FELDMAN, D. A., COLE, N. E., and BAYARD, J. "Prognostic Criteria in Case Histories." *Journal of Consulting Psychology*, 1953, 17, 163-171.
- (2) SCHOFIELD, W., HATHAWAY, S. R., HASTINGS, D. W., and BELL, D. "Prognostic Factors in Schizophrenia." *Journal of Consulting Psychology*, 1954, 18, 155-166.
- (3) ZUBIN, J. "Role of Prognostic Indicators in the Evaluation of Therapy." *Psychopharmacology*, 1959, 583, 343-354.
- (4) LASKY, J. J., KLETT, C. J., CAFFEY, E. M., JR., BENNETT, J. L., ROSENBLUM, M. P., and HOLLISTER, L. E. "Drug Treatment of Schizophrenic Patients." *Diseases of the Nervous System*, 1962, 23, 698-706.
- (5) SHERMAN, L. J., MOSELEY, E. C., GING, ROSALIE, and BOOKBINDER, L. J. "Prognosis in Schizophrenia." *Archives of General Psychiatry*, 1964, 10, 123-130.
- (6) CAFFEY, E. M., JR., DIAMOND, L. S., FRANK, T. V., GRASBERGER, J. C., HERMAN, L., KLETT, C. J., and ROTHSTEIN, C. "Discontinuation or Reduction of Chemotherapy in Chronic Schizophrenics." *Journal of Chronic Diseases*, 1964, 17, 347-358.
- (7) KLETT, C. J., and MOSELEY, E. C. "The Right Drug for the Right Patient." *Journal of Consulting Psychology*, 1965, 29, 546-551.
- (8) NATIONAL INSTITUTE OF HEALTH-PSYCHOPHARMACOLOGY SERVICE CENTER COLLABORATIVE STUDY GROUP. "Phenothiazine Treatment in Acute Schizophrenia." *Archives of General Psychiatry*, 1964, 10, 246-261.

(In the discussion session of the conference, Dr. Klett made these additional remarks:)

It would be difficult at this point in the conference to say anything new about classification. It becomes more a matter of

agreeing or disagreeing with what others have already said, or simply stating an additional opinion. I believe that although classification of patients into types is often convenient for research purposes, it is often unnecessary or even undesirable. It depends upon the research question or upon the goals of the research.

At the same time, I think that the evaluation of drugs, which is the business that I am in, naturally leads investigators to think in terms of typologies, and I would like to give a few examples to illustrate this. In a study of the effects of discontinuing drugs in a group of schizophrenic patients, we defined a group that showed a clinical relapse within a certain period of time and another group that did not relapse within a fixed-time period. We then used discriminate analysis with a variety of rating scale and prognostic variables as predictors or discriminators to see if we could differentiate groups. Now, in approaching the problem this way, we used a type concept of relapsers versus non-relapsers, the purpose being to predict those patients that would relapse. However, since the probability of relapse seems to be a function of time after discontinuation of drugs, we could have equally well approached the problem as one of predicting time to relapse. I don't really think there is such a thing as a relapser type. I think that patients differ in terms of the probability that they will relapse after discontinuation of drugs. However, you don't have to believe in the existence of a type in order to use the concept.

In another study in which we followed schizophrenics for 3 years after admission, we defined a favorable and unfavorable outcome group and used a similar statistical approach. We could equally well have predicted days out of the hospital during these 3 years, or some other outcome variable. Similarly, we once defined over- and under-responders to treatment in terms of the difference between observed rating scale scores after treatment and their predicted

score yielded by a multiple-regression equation applied to pretreatment scores.

In the paper that was circulated, I described an approach we took several years ago in finding the right drug for the right patient. What I meant this to illustrate was that it was not until after the analysis was completed and patients assigned to their drug of choice that there was any thinking in terms of types. Although there was no need to invoke a typology to apply the method, it was felt necessary to characterize the patients assigned to one drug rather than the other. At this point we suddenly had a typology. However, we really had a continuum of scores which were used to assign patients to drugs.

With this introduction, I should comment on the typologies that are described in the three papers. If scores yielded by the Elgin Prognostic Scale or the Phillips Premorbid Scale are used simply to predict outcome of treatment, it is not necessary to think in terms of process or reactive schizophrenics nor is it necessary to use cutting scores except as a convenience, but Dr. Garnezy's excellent review of the history of this concept reveals that it is not simply a device to make life simpler for the researcher or the clinician. It is more in the nature of a personality construct which may have important implications for the underlying biology or developmental history and so on. In this sense it is more of a real typology than the ones I have described. In my paper I commented that although these scales do provide an objective and quantitative operational definition of the types, I was not altogether happy with the scales. For nearly 10 years now I have been, in a sort of trial-and-error fashion, subdividing patient samples on the basis of the background variables said to be associated with prognosis. We have done this using the variables singly and in statistical combinations and have used both the Elgin and the Phillips scales. These efforts have not been very fruitful, and my comment here simply reflects my

impatience that there is not some better approach to the prognostic problem.

I don't want to do Drs. Klein's or Clyde's work injustice by dealing with it too lightly. In some ways it seems to me that Dr. Klein's approach to the right drug for the right patient is essentially the same as the one I described in my paper except that he used clinical insight and observation instead of regression equations and characterizes his groups in quasi-diagnostic terms rather than in terms of the rating scale variables with the largest regression coefficients. As a result, he comes out with something that sounds more like types to me than our results necessarily do. I still have my doubts, however, about how easy it will be for someone else to apply his types because of definitional problems.

Finally, I am so perplexed by Dr. Clyde's type, which is based upon disagreement between doctor and nurse on improvement, that I really hesitate to comment at all. His definition is objective, empirical, and may even have some utility. I said earlier that you don't have to believe in types to use them, and I think also it is not necessary that the types be rational! [Laughter.]

A final few words on my personal experience with typologies. The examples I gave in my beginning remarks could be described as primitive typologies or arbitrary typologies that were either convenient for statistical purposes or for ease in communication. We have also used more seriously proposed typologies. Dr. Anthony Sainz has described two subtypes of chronic schizophrenia that he called hypodynamic and hyperdynamic, and has presented data to show that they responded differentially to drugs, that is, hypodynamic patients responded to drug A but not B, and hyperdynamic patients responded to drug B but not A. This is the sort of problem we are interested in, so we asked Dr. Sainz to help us in defining the groups and in training our clinicians to make the distinction reliably. However, we were not able to replicate his findings with drugs.

More recently, I have been involved with Drs. Lorr and McNair in the identification of naturally occurring groups or patient types. We have now reached the point where we are testing differences between drugs within types. This is not too easily done because not all patients can be classified, and some of the types have a low-base rate of occurrence. Each type has its own set of defining symptoms so it is not satisfactory to use the same criterion of improvement for each type and it may not be possible to make comparisons across types. However, we recognize that this is a critical step in establishing the validity of the types because obviously a typology is of limited value if it does not relate to anything outside the system.

OPEN DISCUSSION

At the conclusion of Dr. Klett's remarks, the meeting was opened for discussion from the floor. Dr. Lubin served as moderator.

Dr. GLUECK. It seems to me that classification is an attempt to order a body of knowledge—maybe in this case a body of ignorance—to group things or people with similar if not identical properties or traits. Presumably if we increase the accuracy of the classification, we increase its power to convey information.

A number of solutions have been proposed to this complex problem. Dr. Zubin, for example, would tackle personality and psychopathology separately. This is, I think, a very dangerous point of view to take.

For years we have been trying to say we need to look at the pathological process occurring within a person and that the concept of disease is the vector of the interaction between the pathological process and the host organism, whether this is bacteria, an environmental insult, or something else.

There are typhoid carriers who carry a bacillus that is virulent to large numbers of people but who get along very happily

with it and show no evidence of pathology, and so on down the line.

Fragmentation is a serious danger of this era of the increasing use of very powerful modifiers of behavior, if not personality. Those who are opposed to anything but the psychogenic theory of the development of psychological illness say, "Use your drugs, but let's get it over with in a hurry so I can really get busy treating this patient with psychotherapy." Or the reverse school of thought says, "Forget about the need to talk to the patient and just find the right drug and give it to him and all the problems will be solved."

If we say psychopathology is one thing and personality is another, we promote this kind of split. I think this would be a most unfortunate result of this conference or our deliberations.

The whole issue of the adequacy or inadequacy of our data that Dr. Lehmann raised and other people commented upon is a very pertinent one, but I don't think it is quite as serious as he has indicated.

Not all clinicians are as expert clinicians as Dr. Lehmann. Two years ago he commented at the College of Neuropsychopharmacology meeting that if you gave him two experienced clinicians—and I stress "experienced"—he would stack them against all the rating skills of the computers and they would win. He's probably right, if he can find two people like himself. But they don't occur very often, and we are forced to deal with much less competent clinicians in the many hospitals and clinics around the world who don't have this kind of clinical know-how. These are the people toward whom much of our deliberations, I think, are directed.

Can we give them some help?

The fact of the matter is, whether we like it or not, we are dependent upon not only the inspectional data (which I think we can improve upon by a number of techniques being developed) but also on introspective data. If the monkeys used in research could only tell us what they were

feeling or how they were reacting, we would be nine leagues ahead. This is what our patients can in fact do and should do. But the problem of how we handle this sort of introspective information gives the statistician the greatest stomach-ache. Inevitably there is some sort of middleman translation, usually the clinician, who distorts the data in unpredictable ways.

Another subject is also germane: As we identify specific etiologies in psychiatry, for example, phenylketonuria, they are no longer our province. Someone else takes over. So that we are left with diseases that are ideopathic or functional, which says simply we don't know what causes them.

I would like to propose an alternative approach: In the clinical laboratory area now with the development of automation, they are discovering here at the NIH laboratory, at the Mayo Clinic, and at other places that it is far more economical from every standpoint to run every conceivable test on a sample of blood, urine, or what-not, and simply present the clinician with all of these and let him pick the ones useful to him. It costs less in the long run. It seems to me we are approaching the same point in the availability of computers and all kinds of information. The information input is tremendous, and we need not only a current cross section but all of the other variables, such as the developmental environment and so on. We could have a multiple-statistical program, with a whole range of statistical analyses, and pick and choose the ones that seem to fit the use of the data at the moment.

Finally, I would urge all of us to remember that, in the long run, our efforts here have to be directed to improving the care of the patient. Unless we do this, there is no sense in our pursuing our deliberations.

Dr. KLEIN. Dr. Klett's skepticism as to the communicability of the types that I pointed out in my paper is shared by practically everybody I know, including myself. All I can say is that, having accumulated suffi-

cient rating scale and objective data on these people, I will attempt shortly to bring out some statistical indices that hopefully will make them identifiable to others.

His paper, if I may discuss the discussant for a moment, is most interesting. I was a little startled at his choice of the criteria of outcome, which was the position of the patient after treatment. It seemed to me that it irreducibly confounds the patients who got to that position because the drug got them there and the patients who were there in the first place and haven't moved.

It seems unlikely that any combination of initial scores is going to make much sense out of such a complex posttreatment situation. Nonetheless, he did get a replication and that in itself is extraordinary. I would like to hear from Dr. Klett as to how the curvilinear bivariate relationships of his prior behavioral scales and their rather odd distributions (J-shaped, etc.) affects the multivariate statistics he applied.

Dr. Garmezy, it seems to me, has obscured an issue by his emphasis on continuum. You may very well have a continuum that is derived from two quite discrete underlying populations.

For instance, let's say we are studying skin pigmentation. Skin pigmentation is largely due to the density of melanocytes in the skin, that can be expressed as a continuous function. Nonetheless, it is quite obvious that we have a multimodal distribution of skin pigmentation. The fact there is a continuum does not speak to the point that there is multimodality and therefore probably qualitatively different ways in which this multimodality came about.

In Dr. Garmezy's paper he points to figure 5 as a collection of process-reactive scores demonstrating unimodality in contrast to Wittman's original report of bimodality.

However, I would submit that his data is clearly bimodal, thus substantiating Wittman, and again raising the issue of a quali-

tative distinction between process and reactive schizophrenics.

Dr. CLYDE. In the experiment I reported in my prepared talk, we studied a certain range of psychopathology over a narrow range of drugs. All we discovered by using elaborate statistical techniques was that something other than chance was influencing the results. But if we want to know how a broader range of symptoms will respond to a broader spectrum of drugs, the experiment does not tell us.

I'd like to see a variety of psychiatric patients treated with a variety of drugs in one systematic study. The data would help us to develop a chart, analogous to the periodic table of elements in chemistry, that would show what symptoms respond to what classes of drugs. Many of the studies reported during this meeting have revealed portions of such a chart, but we cannot yet see the total picture.

Dr. KLETT. I agree with Dr. Klein that it would be more desirable if we had some other criterion of outcome, outside the rating scale system. We worked with what we had.

However, Dr. Goldberg and others at the Psychopharmacology Service Center have generated regressive equations in the same manner I described using a global outcome measure of response to drugs. They also got fairly satisfactory prediction.

Dr. GARMEZY. Regarding Dr. Klein's comments of figure 5 in contrast to figure 2: Figure 2 contains the distribution noted by Dr. Phyllis Wittman in her original study of Elgin State Hospital cases. Figure 5 is based on a very biased sample—the experimental subjects we have used in our research. Since this selection initially was based upon the effort to secure dichotomous groups, we should not be surprised at the presence of bimodality. What is far more surprising is the relatively continuous nature of the distribution of Phillips Scores when you combine the data for the

differing subgroups of schizophrenic patients. I simply don't know what the nature of the distribution of Phillips Scores for a randomly selected sample of schizophrenic patients would be. If drawn from a State hospital population, such a distribution would obviously contain a heavy bias toward larger (i.e., process-type or poor-premorbid) scores.

Finally, with regard to figure 5, I hope that we will not lose sight of my purpose in presenting it; namely, to show that a split by diagnostic categories that was based solely on this subject component may not have fruitful experimental consequences.

There is a further point to be made about multimodality and this transcends the issue of schizophrenia. In the last section of my paper, I suggested that the process-reactive concept may have a generality far beyond its traditional usage—that, perhaps, it could be extended to other psychopathological groups, to children and to so-called normal adults as well. In the closing paragraphs of my paper I present a shortcoming in the design of the Duke experiments—a shortcoming, I may add, present in other investigations as well. An adequate and an inadequate group of schizophrenic patients were compared with each other as well as against an adequate normal group. Clearly what is called for is a control for the inadequate patient group through the addition of process type normal subjects. I hold to the belief that the process-reactive dimension, divorced from its history, is not a specific concept relegated solely to schizophrenia but has a more general applicability. Thus many relationships found to hold in experiments with process and reactive schizophrenic patients may be equally applicable to other patient groups reflecting differing forms of psychopathology as well.

Dr. GRUENBERG. Adolf Meyer once made the witty remark that to use prognosis as a leading criterion in identifying a mental disorder is about as wise as to bring religion

into a political election. Obviously it is not good to bring prognosis into the identification of a condition which you wish to study, particularly if you wish to study outcome. You cannot really investigate the outcome of the so-called schizophrenic patients in our mental hospitals if, in fact, those who make the diagnosis change the diagnosis after a month or two if a patient doesn't do well or if he unexpectedly does do well.

Clearly, it is dangerous to use outcome as an identifying criterion when selecting a group for an investigation of outcome, but I think we do it all the time. Kraepelin did it all his life. He always had some struggles with it, as he became aware that his notions of what determines outcome were incorrect. We have had some good discussions about dangers when selecting cases. I want to point out that we also always select the situation in which we pick our cases. The one thing we fail to study is the movement of the cases into and out of the pool from which selection is made for investigation. How do the cases come to your attention? Which cases haven't been considered for the study?

Also, in most instances, you have not studied the situation under which the patients receive the drugs.

Dr. Garnezy said a few minutes ago that he has begun to think that the problem of process or of deteriorating forces is more general than just schizophrenia. If we ask, "Why do people deteriorate in their personal and social functioning in a hospital?", we get a different perspective on this problem than Kraepelin had. I don't know whether we are right or not, but if we say deterioration might be sociogenic—not psychogenic, not organic—but sociogenic, that it is dependent on the interrelationships of a patient and his total environments, including the institution, the human relationship environment, then one gets the concept of institutional neurosis, as Barton called it, or as some people have called it, a social breakdown

syndrome, secondary to many mental disorders, depending on the conditions under which patients are cared for and their reactions to the environment.

When we studied the problem this way, we found that the majority of cases who deteriorated in the middle age groups, in a community where the effects of institutions are minimized, were not schizophrenic at the time they initially came to clinical attention. They derived from a wide group of clinical conditions; less than half were schizophrenics. The others were a miscellaneous group of other psychiatric conditions.

If you stick to the framework problems, you must also consider the whole question of where are you doing your studies and the general condition of the patient.

Dr. COLE. Dr. Klett raised the predictive problem we've been working with in drug studies, and I think it is relevant to Dr. Klein's point also, as to our population and our drugs.

We have been using a multiple-regression approach—Dr. Goldberg has, and I am speaking for him—and what we find is, given four drugs, four phenothiazines and a placebo, two of the phenothiazines being studied in two separate studies in two independent groups, we can predict response differentially to the four drugs. The predictors check out well back and forth for the two drugs we have studied twice, so it looks good. We plan to replicate it a third time.

Nevertheless, we have a relatively restricted group of patients and a relatively restricted group of drugs; the thing that impresses me is that of something on the order of 36 pretreatment scores on the Burdock-Ward Behavior Scale and on the IMPS, four of them don't predict anything and about 20-odd predict improvement for all phenothiazines and do not discriminate between drugs. We are left then with seven or eight factor scores which show some differential prediction between drugs. An

occasional factor predicts improvement nicely on one drug. But for others you have one score that would predict improvement on three drugs but not on the fourth, or two drugs and not on the other two. This is a patchwork which I have great difficulty conceptualizing into types. One of the reasons I say this is that the types are very useful to clinicians if they can be clear and meaningful.

I think we can predict clinical improvement, given our restricted range of drugs and patients. If you pick the patients who got the drug they would have been predicted, by these equations, to have done worst on, and compare them with the ones who got the drug they would have been predicted to have done best on, one finds about twice as many patients moderately improved or better if they got their best drug than if they got their worst drug.

Dr. KLERMAN. Dr. Garnezy's argument, it seems to me, has failed to clarify a number of issues. One was the issue mentioned by Dr. Klein, distribution versus dimensionality. One can have a variable, like blood pressure, which is continuous, but on which there are unequal distributions, that is, bimodal distributions.

More significantly, it has not been established, as far as I can tell, that the process-reactive continuum is, in fact, a single dimensional variable. Before the discussion of distribution can even be arrived at, advanced psychometric techniques should be applied to ascertain whether or not it is a single dimension. Then we can argue about the nature of distribution, whether it is a one-dimensional or a K -dimensional space.

The second methodological concern I have is the issue of sampling. Most of the studies I have read in the literature have been cross-sectional studies of patients in the acute or chronic building. It seems to me that this has been the wrong method. Cohort analysis would be much more pertinent: to define a group of newly admitted

people and follow them to see if the characteristics present, initially, do in fact influence their behavior 5 years later. This has been a more fruitful approach, I think, than cross-sectional samplings of people who are in back wards versus the admission building, which represent different residuals of cohorts.

Dr. ZUBIN. It is very interesting to note that the process-reactive dimension which Dr. Garmezy considers to be so fundamental not only in schizophrenia but also in the general domain of psychopathology and even normality, should be based so heavily on the Phillips Scale. An examination of the contents of this scale reveals only one underlying dimension—psychosexual development. Why this single dimension should be so important is very puzzling. A review of prognostic indicators finds that psychosexual adjustment, though important, is only one of several hundred traits which are related to outcome (Zubin, J. et al., "A Biometric Approach to Prognosis in Schizophrenia." In P. H. Hoch and J. Zubin (editors), *Comparative Epidemiology in the Mental Disorders*, New York: Grune and Stratton, 1961, pp. 143–203.)

Of course, whether psychosexual maldevelopment is the cause of illness, or whether it is a reflection of illness, is debatable. I am inclined to believe that probably Freud was not quite correct in thinking it is a cause; it is probably a very sensitive indicator, or a sensitive thermometer. This is the most intimate relation a person has; if he fails in that, why, that probably means he is very severely ill. But I wonder what modern improvements such as oral contraceptives and other birth control methods are going to do to this dimension. I wonder if you can continue to hope that this is going to be as effective in the future as it was in the past, since with the removal of taboos, fears of pregnancy, etc., the inhibitions of psychosexual behavior may be lessened.

Dr. Clyde was provocative, but I think he is limiting himself to a situation where regression or correlation situations occur. If you had an homogeneous group to begin with, who varied only in a given trait, why, then, the more of the trait they had perhaps the better would be the relationship to drug effects. But penicillin works on individuals whether they belong to a group of varying intensities of infection, or not so varied. If it is effective, it will work even in a limited narrow range of infections. But suppose you put together patients suffering from five or six different illnesses, some of which are not affected by penicillin? Here the relationship between drug effects and degree of infection will be difficult to trace. That's the problem we face whenever we do a study, the fact that the patients are not homogeneous, and do not respond linearly to drug effects.

Dr. Glueck made a very important point about the separation of personality from psychopathology. But I didn't mean to say we should separate them actually in the study: I said we should separate them theoretically in concept; that the systematic way a person behaves is his personality. This exists whether he has psychopathology or not, before, and after, and in between. It is this personality, perhaps, which tells us how his illness will progress or how it will be relieved. Therefore I consider it very important. But do not separate the two; have them go hand-in-hand. The danger lies not in considering them, but in confusing them with each other.

Dr. GARMEZY. On the issue of acuteness-chronicity, I have cited, primarily, studies that have used fairly acute cases rather than back-ward samples. Acuteness, as you know, has a confusing definitional quality—it can refer to the dimension of prognostic efficacy or to length of illness and institutionalization. Many of the studies I have described use matched groups of acute

cases in the sense that both goods and poors are early equated either for length of hospital stay or admission status.

As for multidimensionality, one recognizes, of course, that a wide range of attributes distinguishes so-called process from reactive cases. The recognition of that fact must be separated from the questions of experimental strategy and the nature of instruments available for producing a rational and theoretically appropriate separation of cases. I have tried to point out that the Phillips Scale permits a distribution of cases that extends along a continuum of premorbid adequacy. The experimental strategy that is oftentimes used is to test only at the extremes of the dimension. However, for those few studies that

have made use of an intermediate group, there is some minimal evidence that such subjects differ from others drawn from the more extreme points of the distribution.

Several of the issues that have been raised by the discussants could better be resolved if we had available for study remitted groups of process and reactive patients—recognizing the reality that remission rates in the two groups differ markedly. Nevertheless, such studies, akin to those conducted with more acutely disturbed patients, would provide potential insights into those subject attributes that can be best viewed as state variables in contradistinction to those that have a more transitory quality and are present only during the manifest psychosis.

D. Systems Which Attempt to Encompass Variables From Several of the Significant Dimensions of Functioning in Mental Illness

Typology of Schizophrenia Based on Multidisciplinary Observational Vectors¹

Nils B. Mattsson, LL.M.² and Ralph W. Gerard, M.D.³

The need for isolating meaningful types of schizophrenia has been obvious long before Kraepelin and Bleuler made their contributions to the understanding of the illness. Objective typologies or classifications of schizophrenia are of utmost importance as regards questions about etiology, treatment, and prognosis. In this respect, the validity of the psychiatric diagnosis has not improved much since the turn of the century. The problem is bad enough in the ordinary fields of medical diagnosis, where there are concrete indicators, as Koplik spots in measles or other clear manifesting signs or symptoms. In the area of mental disturbance the problem has, of course, been confounded by the fact that the diagnosis has originally been reached by clinicians in terms of the, mainly, subjective evaluation of behavioral phenomena.

Most of the published typologies of schizophrenia have involved matching of patients' often subjectively evaluated characteristics within one distinct area of interest, such as psychiatric ratings of pathology, socioeconomic background data,

or drug response patterns stemming from some assessments of improvement or change. No single classification can be exhaustive by itself because subjects who are alike on one set of characteristics may not be similar on another set. Also, the lapse of time may change the profile of characteristics to the extent that what seemed to be a well-defined type at time t_i is a mixture of dissimilar people at time t_{i+k} . Or, looking at the problem from another angle, people in a time t_i cluster may not display the same change curves over the intervals t_{i+1} , t_{i+2} , . . . , t_{i+k} . An attempt to type schizophrenics on response to drugs over time has been made by Rice and Mattsson (in press).

The studies described here involved more or less simultaneous observations or objective tests covering the areas of social history, ward behavior, psychiatric ratings, psychological tests, anthropometric measurements, physiological responses, and biochemical or bioassay tests on body fluids. The observational vector for each of the 208 subjects⁴ in the final stage of the study (pilot and control studies excluded) contained 388 variables which were fairly evenly distributed between the behavioral and biological areas, and were either known or believed to relate to schizophrenia in one way or other. The patients, all male,

¹ The research described in this paper was done under the auspices of the University of Michigan (Department of Psychiatry) and Ypsilanti State Hospital over a period of 6 years (1957-63), and supported by USPHS grants MY-1971 and MY-1972.

² Biometric Laboratory, the George Washington University, Washington, D.C. Now with Advanced Systems Development Division, IBM Corp.

³ Dean of the Graduate Division, University of California, Irvine, Calif.

⁴ Actually over 650 patients, prisoners, and "normals" were tested at one stage or another of the study.

of whom about half were schizophrenic and half different types of nonschizophrenic were kept under controlled conditions, including diet, on a research ward. Prior to testing and evaluations the patients had been off medication for a minimum of 4 weeks. The age range was 18–50, and maximum length of hospitalization 10 years. Fifty percent of the schizophrenics had been in hospital less than a year, while this was true for 65 percent of the nonschizophrenics. The diagnosis was based on the demand that three psychiatrists independently concurred in calling a patient either schizophrenic or nonschizophrenic. Of the 107 schizophrenic patients 42 carried a research diagnosis of paranoid reaction, while most of the other 65 were called chronic undifferentiated. The background, procedures, and selection of patients and test batteries have been described in more detail in: R. W. Gerard. "The Nosology of Schizophrenia: A Cooperative Study." *Beh. Science*, 9, 4, 1964.

Theoretical and Methodological Aspects

A statistical typology, to be complete, is composed of at least three distinct parts. Kossack (1963) has suggested the term "discernment theory" for this kind of analysis.

1. The first part of the analysis starts with a sample of members which may or may not emanate from the same population. The p -variate observational vector is obtained for each member. On this basis a decision has to be made as to whether or not the members fall into clusters. This is what Kossack calls the statistical sorting or numerical taxonomy problem. While on one hand there are no mathematically proved methods available for the multivariate case, there are, on the other hand, quite a few so-called clustering techniques the basic ingredient of which is some kind of similarity index based on profile shape and/or distance. In publications distributed to the participants of this conference,

Zubin (1964) discusses some useful techniques and Saunders (1964) gives a method of his own, adding it to a class of statistical tools he calls syndrome analysis. We would like to call the reader's attention to three additional techniques for solving the clustering problem.

Kendall (1965), who calls this the classification problem, has developed a simple method, the order statistic method, which is distribution free and can be applied to data which are measurable or polytomized or are a mixture of both. Kendall claims that his method works "reasonably well." It builds on an index S which can be regarded as a kind of distance between the members of a pair. The index is given by:

$$S = \frac{p}{\sum_{i=1}^p} \frac{(X_i - X_i')^2}{\text{var. } X_i}$$

where p is the number of variables, X_i and X_i' are the values of the i th ranking for the pair in question and the variance X_i is calculated as

$$\text{Var. } X_i = \frac{1}{12n} ((n^3 - n) + \sum (t^3 - t))$$

where the summation is over all ties of extent t and n is the number of ranks (Kendall, 1958). The clustering starts with the pair closest together. Then each additional member has to be chosen so that the mean distance within the cluster increases the least. When a criterion of largest permissible mean distance has been met, a new cluster is started by choosing a new pair of close individuals.

Another distribution-free approach to this problem has been developed by Tanimoto (1960). Each variable is considered to be of a simple attribute (1 or 0) type; e.g., male or not male, severely ill or not severely ill, 15–20 years old or not, etc. The clusters are obtained through special (ranking) operations on a matrix of similarity indices based on the ratio of matches (1 and 1) to nonmatches (0 and 1 or 0 and 0) over the array of attributes for each pair of subjects.

The third method called the successive factorial screening technique was developed by the senior author (Mattsson, 1963) and used to cluster the subjects in the studies discussed here. The measure of agreement for a pair of subjects can be a product moment correlation coefficient, a coefficient of congruence (Tucker, 1951), or any other cross-product index that can be submitted to principal components analysis. Because the agreement index is calculated between people over the array of measures, all scores have to be standardized to insure equal metric. This method resembles Tanimoto's analysis to the extent that clusters are removed from the matrix as they are found, the successive operations being on members not yet removed from the matrix. It builds on the fact that the first principal component is the best single dimension to summarize the shared variance expressed by the matrix of agreement measures; e.g., a correlation matrix. It involves a successive extraction of the first (and largest) principal component of the matrix, each time reduced by the rows and columns corresponding to subjects meeting the criterion of belonging to a cluster. The operation is terminated when no more clusters can be formed, or when there are too few subjects left for a definable cluster. The criterion of belonging is a smallest acceptable loading (projection) on the principal component. If the extracted dimension is bipolar, only the pole with the larger number of clustering subjects is chosen; the subjects on the other pole remain in the matrix. When all possible clusters have been extracted, the centroid for each cluster may be computed, the correlation coefficients between cluster profiles and subject profiles obtained, and each subject assigned to the cluster with which he has the highest correlation. This procedure will correct for some initial misclassification provided that the clusters are reasonably large to give stable estimates of centroids. This method will not yield in-

dependent types of people; they will be more or less correlated. We do not believe, however, that orthogonality is a basic requirement for a typology in this area. This technique bears a close resemblance to the method of *Q*-factor analysis where people instead of tests are factored.⁵ In this process all information about underlying characteristics disappears and the dimensions have to be defined in terms of subjects. This makes it necessary to refer back to the common characteristics in describing the types. The method of *Q*-factor analysis was also used in our studies.

2. The second stage of a typology is, using Kossack's terminology, the identification or discrimination problem. Given the clusters from the sorting stage, assuming they represent some distinct populations of interest, and for whom a *p*-variate observational variable (either the same one as used in the sorting problem or another relevant one) is available, the identification stage involves finding a numerical measure that best discriminates between the given populations. This is usually accompanied by a test of significance for the group separation in the multidimensional space. If significant, it makes the investigator comfortable in regard to his endeavor to produce a meaningful typology. A scrutiny of the significant dimensions making up the space may also aid in the interpretation of the types. The best known solution to this problem is obtained by the use of a linear discriminant function.

3. The third part of a typology is the classification or diagnosis problem. This has to do with the assigning of a new or additional individual to one of the given populations assuming that he came from one of them and knowing his multivariate observational vector. To achieve this goal a method which usually is called a decision rule has to be constructed. This often in-

⁵ The transpose factoring technique was devised by Burt (1937) and Stephenson (1936). Additional discussion of this much debated technique can be found; e.g., in Thomson (1950), Cattell (1952), and Harman (1960).

volves the reduction of the multivariate problem into a univariate problem by finding the best linear combination of the original variables. In a two-group case one discriminant function is sufficient to classify an individual. The rule to be set up states that if the individual's "score" on the function is larger than a given cutoff value he belongs to population A, otherwise to B. There are considerations about misclassification, nonlinearity of the function, etc., which, although not discussed here, are of vital importance. The discriminant function can be extended to cover k groups, in which case usually several functions have to be considered in making decisions about an individual. In the most parsimonious case one function may be sufficient which means that all groups lie in one-dimensional space. Cooley and Lohnes (1962) have published a computer program which classifies individuals according to the most probable of k (k =number of populations) selection hypotheses. The basic concept is the centroid score which is used to assess "the degree to which an individual resembles each of several groups in terms of certain antecedent variables that are known or believed to be important in distinguishing among these several groups." The input to this program consists of the results from a discriminant function analysis. There are several other methods available for assigning an individual to his proper population when the population parameters are known or estimated from samples. The one given by Rao (1952) and called the linear discriminant score was used in the studies under discussion to check possible misclassifications and to give a rule for classification of new individuals. The linear discriminant score corrected for a priori probabilities is calculated for each sample and the individual is assigned to the population for which his score is highest. If the probabilities are not known, as was assumed in our study, the maximum likelihood method leads to the rule of assigning an individual to that population

for which his uncorrected discriminant score is highest.

Typologies

The idea of breaking up the clinical entity schizophrenia into two or more types was based on the hypothesis that schizophrenia is a nosological mixture and that the separate components could be indicated and identified by analyzing profile similarities and dissimilarities between a sufficient number of patients on a wide variety of tests. Because the correlations of the tests with the criterion, the typology sought, were necessarily unknown, the tests had to be selected on some other basis, such as relevance to schizophrenia. Most tests or observations are not sharp enough for this purpose; i.e., to find the fine details of behavior, be it psychological test performance or perhaps some response due to biochemical activity. From a universe of possible measures it may prove very hard to find a sample of really relevant ones. Our criterion is the psychiatric diagnosis, the very one that we try to ignore later on.

From the 388 test or observation variables selected for the main study, 180 were chosen for the typology problem. This selection was prompted by the need to rid the matrix of excess redundancy and to bring balance to the representation of behavioral and biological areas. Of the 208 subjects in the main study we could use only 147 in this analysis, because the rest did not have complete data on all variables included. By clinical diagnosis there were 80 schizophrenics and 67 nonschizophrenics. The schizophrenic group was simply divided into paranoid and nonparanoid; and the nonschizophrenic group into those with chronic brain syndrome, and others.

In this presentation the results of two major approaches to the classification problem will be discussed. The first one was a factor analytic approach involving both schizophrenic and nonschizophrenic patients, while the second approach was

through clustering techniques and involved only schizophrenics. If there exists a combination of some ideal measures, which correlates perfectly with some natural typology of schizophrenia, we do not claim that we have found it, despite the wide field covered by the tests and observations in our study. We can, however, present two multidisciplinary typologies of schizophrenia in the hope that they will add to the pool of knowledge.

1. Q-Factor Approach

In our first analysis we decided to include all subjects, regardless of diagnosis, because we thought that, in addition to direct misdiagnosis, there must be a number of non-schizophrenic mental patients in whom traits, commonly considered characteristic of schizophrenia, are present to various degrees. On the basis of this reasoning, we should expect to find, in addition to more or less clear types of schizophrenics and nonschizophrenics, one or more mixed types. Conversely, from the results of such a typology one might be able to reason out the answer to the question of what is purely schizophrenic behavior or response and what is common to both groups being thus confounded and hard to attribute to either one specifically.

Before computing the correlation matrix between the patients all scores were standardized to equal metric. The statistical sorting was based on a principal components analysis (Harman, 1960) of the correlation matrix. What is called similarity of a pair of patients is based completely on the shape of their profiles; the distance between them is ignored. If distance had been accounted for in addition to shape, more clusters would probably have resulted. Eleven significant factors⁶ were extracted and, after inspection, rotated orthogonally. Actually, two rotations were performed, one using the normal varimax criterion

(Kaiser, 1958) and the other the quartimax criterion (Neuhaus and Wrigley, 1954). The latter, producing a clearer pattern, was adopted.

All type factors were bipolar with sufficiently high (patient) loadings on both poles. The types could consistently be divided into two groups: The clinically less disturbed and the clinically more disturbed. Together they represent a type characterized by some common features, but with opposite signs. One pole of a factor seems to be needed to bring out the other, usually the one with the clinically more pathologic patients. The group manifesting less clinical pathology will consistently be called the positive group. Whenever such words as "less" and "more" are used without a reference point, the comparison is with the mean of all patients in the sample.

Taking a cutoff loading of ± 0.30 to indicate who belongs to a type factor, about 25 percent of the patients overlap with another factor (split loadings). This is the same as saying that these patients belong to two types, which could be explained as, for instance, a behavioral-biological split. Equivalent to split loadings is the clinical practice of making split diagnoses when the patient cannot clearly be assigned to one category only; e.g., schizophrenic reaction, catatonic type with paranoid features. Naturally, we could have made a forced classification, assigning a patient to the type with which he had the higher correlation, or assigning him to the unclassified. In the description of types, the overlapping patients have been included in view of the fact that the factors are orthogonal and a patient's contribution to one factor is thus independent of the contribution to another.

The identification (discrimination) of these types was not accomplished by any multivariate statistical approach, mainly because it would have been absurd to run a multiple discriminant function involving 180 variables and then give the combina-

⁶ All roots equal to or larger than 2.0 were considered to be significant in view of the large correlation matrix (order 147), instead of the usual criterion of roots ≥ 1 .

tions of these as identifying equations. Subsets of variables were tried out but did not, in general, reproduce the types. Instead, the characteristics underlying the type factors were determined from the correlations between these and the test vectors.

Q-Type 1

	N	Mean age	Mean hospital stay, months
+ group.....	33	34	4
- group.....	21	37	40

The first type factor contains a mixture of schizophrenics and nonschizophrenics but the clinical types are rather clearly separated from each other. Highest loading on the positive pole are nonschizophrenic personality disorders, followed by slightly lower loading paranoid schizophrenics. A group of nonparanoid schizophrenics loads highest on the negative pole, followed by a mixture of brain-damaged patients and nonparanoid schizophrenics. Factor one is mainly a dimension of psychological functioning and general psychopathology in that the positive end is represented by patients who are relatively well functioning on psychological tests, come from higher social class, and show relatively little disturbance on psychiatric and ward behavior ratings. The patients representing the negative end are psychologically poorly functioning, come from lower social class, and manifest a high degree of overall psychopathology (conceptual disorganization, motor disturbances, perceptual distortion, withdrawal, etc.).

Q-Type 2

	N	Mean age	Mean hospital stay, months
+ group.....	12	39	16
- group.....	11	28	58

This type is best characterized in terms of emotional reactivity, social adjustment,

body-build, and orthosympathetic activity. The positive group consists of patients with mixed diagnoses (nonschizophrenic, chronic brain syndrome, and paranoid schizophrenic) while the negative group contains mostly nonparanoid schizophrenic. The positive group is relatively free of anxiety, had good peer social adjustment before hospitalization, has a generally heavy body-build, and shows relatively high orthosympathetic activity (blood pressure increase in response to epinephrine and norepinephrine, increased excretion of epinephrine and norepinephrine especially in relation to 5-hydroxyindoleacetic acid). The negative group is highly anxious and self-depreciative, was socially withdrawn in childhood and adolescence, has a small and light body-build, and shows lower than average orthosympathetic activity.

Q-Type 3

	N	Mean age	Mean hospital stay, months
+ group.....	9	26	20
- group.....	9	42	49

The positive group of Q-type 3 contains patients with mixed diagnoses. The negative group is composed of both brain-damaged and other nonschizophrenic patients with alcoholism. The discriminating features of this type are withdrawal, importance of role of patient's father, parasympathetic activity, and general autonomic response to psychological stress. The positive group, characterized by less withdrawal, comes from highly disorganized, conflict-ridden families with irresponsible fathers. These patients were, however, highly motivated in school. Their autonomic response to psychological stress (shock, tone, etc.) is below the mean of other patients. The negative group, characterized by alcoholism, comes from stable families with conscientious and occupationally and socially successful fathers.

These patients, although relatively intelligent, show memory impairment, and are highly withdrawn with increased parasympathetic activity (ratio of blood pressure change to heart rate change on norepinephrine). They also show increased autonomic reactivity to psychological stress. A significant additional feature of the negative group is the low urine excretion, on the average less than a liter per day.

Q-Type 4

	N	Mean age	Mean hospital stay, months
+ group	7	29	10
- group	5	45	78

Type 4 resembles Q-type 1 in that it is mainly characterized in terms of psychological functioning, although determined more by psychological test behavior than by psychiatric and ward observations. The positive group consists of patients with mixed diagnoses, mostly nonschizophrenic; the negative group has chronic brain syndromes and nonparanoid schizophrenics. The group on the positive end functions well on psychological tests, whereas the performance of the negative group is considerably below the average, especially on intellectual and psychomotor speed tests. The latter group also has an unusually high urine volume, on the average almost 4 liters for a 24-hour period. The night urine volume alone is 1,200 cc., or close to what a normal person excretes in 24 hours. Deviant glucose tolerance is another significant characteristic of Q-type 4; three of the five patients in the negative group and two of the seven in the positive group have diabetic tolerance curves.

Q-Type 5

	N	Mean age	Mean hospital stay, months
+ group	8	35	61
- group	7	41	55

The positive group has patients with mixed diagnoses while the negative group is clearly schizophrenic (paranoid and non-paranoid). Q-type 5 is best described by the negative group which gives a clear picture of a classical textbook schizophrenic: High intelligence; low motivation; slowing down of psychomotor response; high degree of mental disorganization; and skinny and tall body-build. Peculiar to this group is also a very low glucose absorption and utilization rate, and low autonomic reactivity. The clinically relatively little disturbed positive group, consisting of patients with below average intelligence, does not contribute much to the description of Q-type 5.

Q-Type 6

	N	Mean age	Mean hospital stay, months
+ group	10	35	14
- group	5	41	65

Q-type 6 is best described in terms of psychopathology and autonomic response to psychological stress. In the positive group are patients with mixed diagnoses; the negative group is again clearly schizophrenic. The positive group is among the most active, the negative (schizophrenic) group among the most underactive and withdrawn of all patients. This underactivity is reflected even in their autonomic response, revealed especially by a reduction in heart rate and respiration rate under psychological stress. On the biochemical side the negative group has a high BSP retention, over 8 percent, compared to the normal upper limit of 5 percent, and the positive group a relatively high creatinine excretion. It seems that the inactive schizophrenic group "grew up in a sitting position": The trunk is long, the pelvis is wide, the armspan is long, and the legs are short.

Q-Type 7

	N	Mean age	Mean hospital stay, months
+ group.....	7	41	16
- group.....	11	29	17

Both groups in Q-type 7 contain mostly schizophrenic patients. The clinically less pathologic positive group is the most compliant, least resistive, and least paranoid of all patients. Yet, the social background data describe these patients as aggressive and easily frustrated. The negative group patients come from lower class homes with fairly comfortable relationships within the home. Psychiatric ratings describe them as highly paranoid and resistive. The two groups respond in opposite directions under certain autonomic stress. The negative group shows a heart rate increase with epinephrine, whereas the positive group shows a decrease. Similarly, blood pressure and heart rate are reduced for the negative group on psychological stress, with the reverse true for the positive group. On biochemical measures, epinephrine and norepinephrine excretion is high in the positive group, fairly low in the negative. The negative group is also characterized by lower potassium and sodium in urine, and higher blood creatinine.

Q-Type 8

	N	Mean age	Mean hospital stay, months
+ group.....	8	44	13
- group.....	8	26	15

Q-type 8 can be best identified by social history characteristics and circulatory response to autonomic drugs. The positive group, mostly older nonschizophrenics, is clinically little disturbed, well managed and active on the ward. Diagnosed as alcoholics, they come from low social class backgrounds, had ineffective, alcoholic

fathers and overprotective mothers. They were socially passive and withdrawn, and highly dependent on their mothers. The mean blood pressure baselines for the group are low and the heart rate changes very little in response to autonomic drugs. Liver function deficiency is indicated by decreased albumin levels and the albumin/globulin ratio. The negative group contains mostly young schizophrenics rated as very underactive and resistive on the ward. These patients come from higher social class with fathers who were fairly efficient in their role and mothers who were ambivalent toward their sons. Their heart rate response to autonomic drugs, especially yohimbine and piperoxan, is unusually high.

Q-Type 9

	N	Mean age	Mean hospital stay, months
+ group.....	11	28	21
- group.....	7	41	31

The negative group in Q-type 9 looks like a classical textbook case of flat, burnt-out schizophrenia without overt psychotic manifestations. Psychiatric ratings describe them as submissive, underactive, and the most withdrawn of all patients. As children they were passive, underactive, and driven to accomplish in school. Psychological tests show them functioning at a rather high intelligence level with good verbal ability. On psychomotor tests, however, they are very slow. Physiological findings indicate a high degree of general autonomic reactivity in this group. The positive group, mostly nonschizophrenics, presents a rather mixed, unclear picture both psychiatrically and psychologically. Social history scales reveal extremely poor occupational and educational achievement, overactivity, poor frustration tolerance, aggressiveness, and a family background of conflicts. The ward behavior ratings describe them as management problems.

Q-Type 10

	N	Mean age	Mean hospital stay, months
+ group.....	9	34	6
- group.....	6	30	14

The positive group consists of nonschizophrenics, the negative group mainly of schizophrenics. The patients in the positive group, mostly diagnosed as alcoholics, come from unstable, conflict-ridden families with irresponsible, hostile parents. The negative group, rated very anxious and the most paranoid of all patients, brings out the schizophrenogenic mother syndrome: their mothers were closely tied to the home, overpossessive, and overprotective. The groups are about average in their psychological test behavior. The negative group patients, however, seem to be quite unrealistic in that they show marked discrepancy between their level of aspiration and actual achievement. In the autonomic area they have a significant increase in galvanic skin response during the baseline session, and in secondary heart rate change on yohimbine; the positive group shows a decrease on both measures.

Q-Type 11

	N	Mean age	Mean hospital stay, months
+ group.....	7	28	7
- group.....	7	41	23

The identifying features of this type have to do with general psychopathology and temperature and respiration response to psychological stress. The positive group patients are outgoing and well integrated (least mentally disorganized of all patients). Psychological tests characterize them as intelligent and well coordinated. They were active as children and had good tolerance for frustration. The patients in the negative group are underactive, withdrawn, and

mentally disorganized with about average performance on psychological tests. They were less active as children and had a very poor frustration tolerance. Typical of the negative group are especially: A higher than average skin temperature baseline and a lower than average, but disturbed, respiratory rate baseline, lowered skin temperature and increased respiration rate on psychological stress. Also notable is the higher total catecholamine and epinephrine excretion.

As already indicated above, no specification equations are provided on which to classify additional or new individuals. It is, however, possible to assign an individual to a type on a descriptive basis, given that all, or at least most, of the pertinent information is available. In this case only the presence or absence of a characteristic would be taken into account, whereas the number and quality of the fitting attributes required to classify a subject would depend on the classifier.

The described types reflect many popular views about schizophrenia, such as:

1. The paranoid-nonparanoid breakdown displayed by Q-type 1.
2. The classic textbook schizophrenics in Q-types 5 and 9, one anxious and mentally disorganized with no intellectual deficit, the other also intellectually intact but inactive, withdrawn, burnt-out residual.
3. The split between outside and inside behavior (Q-type 9): Schizophrenics inactive on the outside are highly reactive on autonomic nervous system tests. There are, however, also contradicting trends: In Q-type 6, e.g., psychological and autonomic inactivity are positively correlated.
4. The defective carbohydrate metabolism reflected in Q-types 4 and 5.
5. The schizophrenogenic mother syndrome revealed by Q-type 10.

Several additional Q-factor analyses were tried both on the full sample and on

schizophrenics alone, involving all of or different subsets of the 180 variables. A few of them will be briefly discussed here.

Separate analyses of the behavioral and the biological variables did not show many common clustering subjects. The behavioral clusters bore more overall resemblance to the clusters in the combined analysis, although the latter was not reproduced by either one.

An analysis of schizophrenic patients only on all 180 variables replicated fairly well only the first 2 factors in the combined sample. The two groups on the first factor were composed of practically the same well-functioning paranoids and poorly functioning nonparanoids as on the first factor of the combined sample. The rest of the factors contained too few clustering subjects for any safe estimates.

We also decided to summarize all the information obtained on the schizophrenic

patients by scoring them on 34 test factors representing all areas of study. Neither an analysis involving these artificial variates and the schizophrenic patients nor another analysis with 24 objective test variables⁷ did replicate the earlier derived types, except for those on the first, and to some extent, on the sixth factor. A new pattern resulted, however, which seemed to take better shape in a subsequent analysis. The psychological variables still tended to divide all subjects into two major categories of well and poorly functioning, but now it seemed that the paranoid-nonparanoid dimension was the main cause of this separation. The groups on, in each case, six bipolar factors tended to be either one or the other.

The 24 objective test variables were chosen from the areas of psychology (12),

⁷ See table 1. These variables were also used in the clustering techniques described in the next section.

TABLE 1.—Tests Used for Statistical Sorting and for Discriminant Analysis

Variable name	(1) Used for—		Identification number in discriminant analysis
	Sorting	Discriminant analysis	
Psychological:			
Facial expression, male faces.....	(1)		
Facial expressions, female faces.....	(1)	(1)	1
Graded analogies.....	(1)	(1)	2
KTSA, level of thinking.....	(1)	(1)	3
Social concepts.....	(1)		
Wechsler-Bellevue, digit symbol.....	(1)		
Dotting, preferred speed.....	(1)		
Dotting, fastest speed.....	(1)	(1)	4
Reaction time, uniform interval.....	(1)		
Reaction time, irregular interval.....	(1)	(1)	5
Stroop Color-Word Test, score 1.....	(1)		
Stroop Color-Word Test, score 2.....	(1)		
Autonomic nervous system:			
Resting blood pressure, standard deviation.....	(1)		
Resting skin temperature.....	(1)	(1)	6
Resting percent disturbed respiration.....	(1)		
Respiration rate change with shock.....		(1)	7
Heart rate change on epinephrine.....	(1)	(1)	8
Blood pressure change on epinephrine.....		(1)	9
Heart rate change on norepinephrine.....	(1)		
Parasympathetic index.....	(1)	(1)	10
Biochemical:			
Urine volume, 24 hour.....	(1)	(1)	11
Creatinine blood/urine.....	(1)	(1)	12
Sodium (urine), 24 hour.....	(1)	(1)	13
Potassium (urine), 24 hour.....	(1)	(1)	14
Humoller rate.....	(1)		
Serum phosphorus.....	(1)	(1)	15

physiology (6), and biochemistry (6) so that:

- (a) The balance between the behavioral and biological areas would be maintained.
- (b) They significantly discriminated between schizophrenics and nonschizophrenics (there are only few such tests in physiology and biochemistry).
- (c) They would as much as possible represent the stronger test factors in their areas.

It was decided to use this subset of test variables for further development of the paranoid-nonparanoid pattern, mainly because it produced practically the same 6 type factors as those obtained with the 34 factor scores.

2. *Clustering Technique Approach*

In order to (1) clear the types of patient overlap, (2) achieve a possible reduction in the number of types accounting for the patient sample (schizophrenic only), and (3) lose as much as possible of the bipolarity experienced with the *Q*-factors, we agreed on using techniques which sort a subject into one of the several clusters to be formed or leave him unsorted depending on some previously determined criterion. Bipolarity will always be present to some extent, although it is not quite obvious at first glance. In most typologies derived through clustering techniques one can find types which are each other's opposite to a considerable degree, thus de facto lying in one dimension. In general, there will be correlation between types varying from high negative to moderate positive. This is due to initial both positive and negative relations between individuals, regardless of the fact that the underlying variables may have entirely positive intercorrelations.

As Tanimoto's taxonomy analysis assumes only positive or zero relationships between individuals, it seemed a suitable method to try. All variables had to be divided into mutually exclusive class inter-

vals so that a subject's quantitative score on a variable would fall into one category only. To provide equal weight to the variables, all of them were divided into the same number of categories. A trichotomy based on 33 percent areas under curve was chosen. The range of each variable was thus divided into three parts with approximately equal frequencies. This provided one of many ways of cutting the variables, and the resulting subject clusters naturally depend on this decision. The analysis yielded five clusters which did not resemble any of the previously obtained patient types. Discriminatory analysis did not produce significant group separations on the variables used. Few other project variables were significantly related to these groups.

The unsatisfactory results were at first attributed to the fact that the 24 variables had been used instead of the 34 factor scores. After a similar analysis, using the latter, did not come out any better—seven clusters with little resemblance even to the five mentioned above—the method was dropped. It is, however, known to give good results with strictly categorical variables, and also with ordinal data which are much easier to categorize than quantitative data. The method is probably very sensitive to the choice of class intervals with quantitative data.

The application of a clustering method given by Holzinger (1941), utilizing an index of similarity called the *B* coefficient or the coefficient of belonging, was considered next. The coefficient is defined as 100 times the ratio of the average of the intercorrelations among the variables (subjects) of a group to their average correlation with all remaining variables (subjects). It proved, however, difficult to use with a matrix of people correlations, because of the existing negative relationships mentioned above. The coefficient was originally devised to sort out variables into groups using overwhelmingly, or entirely, positive intercorrelation matrices. The negative correlations affect the denominator of the

B ratio so that it takes values from some positive number through zero to some negative number. The ratio accordingly takes values from some positive number to $+\infty$, then jumps to $-\infty$, whereafter it approaches zero, the sign remaining negative. This technique has probably been modified by many users to take into account the negative relationships. Rice and Mattsson (1965), e.g., have modified it and written a Fortran computer program for the IBM 7094.

At this point, the technique called successive factorial screening was introduced and applied to the *Q*-intercorrelation matrix of the 80 schizophrenic patients on the 24 objective test variables. Although it involves the calculation of successive principal components, we think the method belongs with the clustering techniques. The difference is that in clustering methods usually an average involving some agreement indices, forming a tentative cluster, is contrasted with another average involving the indices supposedly not in the cluster, whereas the screening method checks the peoples' correlations with a common, artificial reference vector, the major axis of an ellipsoid, on the basis of which a cluster is determined. The similarity is in that clusters are removed as they are formed and the subsequent operations are on the individuals' agreement indices which are left in the matrix. The screening method is thus only looking for the heaviest concentration of shared variance at each successive operation in contrast to the ordinary principal components analysis, where an orthogonal reference frame is sought.

The criterion of belonging to a cluster was determined as a correlation (loading) better than 0.40 between an individual and the reference vector. The method produced seven clusters of individuals, leaving four subjects ungrouped. Compared with the *Q*-factor analysis of the same subjects on the same 24 variables, there were now 7

clusters (instead of 12 on 6 bipolar factors). The clusters obtained in these two analyses bear a clear resemblance to each other, although some shuffling of subjects has occurred in forcing them into fewer groups.

Three of the seven clusters consisted of patients mostly diagnosed as paranoid schizophrenics, while the majority of the patients in the remaining four clusters were diagnosed as nonparanoid schizophrenics.

For further statistical workup redundant variables which were useful at the sorting stage, were dropped and two variables in the autonomic area were replaced, when a preliminary analysis of variance indicated that a better discrimination of the types could be obtained on this basis. Fifteen variables (table 1) were thus used for the discrimination and the classification problems. These variables are sufficient to identify the types. Univariate analyses of variance indicated, however, significant differences (*F* ratio for main effect corresponds to a $p < 0.05$) between types on practically all project variables, including biological and social history variables which had not shown much differentiation before. A major part of the differentiation was accounted for by the paranoid-nonparanoid split.

The grand means and standard deviations of the variables involved are given in table 2. Based on these values standard scores were calculated for each subject on each variable. The centroids, in standard score units, for the seven types are also given in table 2.

A summary of the numbers and percentages in each type is given in table 3.

A multiple discriminant function analysis, summarized in table 4.1, indicates a significant group separation which can be explained in, essentially, three dimensions. The group centroids in discriminant space are pictured in figures 1, 2, and 3. The actual values are given in table 4.2.

TABLE 2.—Group Means and Grand Mean and Standard Deviation on 15 Variables in Discriminant Analysis of 7 Schizophrenic Types

Variable Number	Explanation	All subjects		Group means in standard score units						
		x ¹	σ	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6	Group 7
1	Facial expressions, female.....	36.32	5.31	0.83	-1.14	-0.32	0.73	0.65	-0.22	-0.31
2	Graded analogies.....	13.87	6.42	.78	-.69	-.50	1.11	-.26	-.51	-.08
3	KTSA, level of thinking.....	67.83	26.98	.88	-.72	.14	.52	-.09	-.30	-.91
4	Dotting, fastest speed.....	55.12	17.86	.73	-.94	-.15	.24	.96	-.73	.00
5	Reaction time, variable delay ²	165.10	27.34	.41	-.70	-.57	.39	.41	.32	.31
6	Resting skin temperature.....	29.54	3.31	.25	-.12	.26	.12	-.57	.63	-.91
7	Respiration rate change with shock.....	.93	2.44	-.05	.13	-.21	-.24	.96	-.40	-.42
8	Heart rate change on epinephrine.....	5.65	9.40	-.28	.12	.23	-.06	.66	-.87	.10
9	Blood pressure change on epinephrine..	24.64	13.45	.31	-.09	-.55	-.25	-.05	1.08	-.33
10	Parasympathetic index.....	-.51	1.62	-.55	.43	-.22	.19	-.55	.83	.20
11	Urine volume, 24 hour.....	1.54	.66	.43	.03	-.97	-.40	1.29	-.14	-.55
12	Creatinine blood/urine.....	.93	.39	-.60	.03	1.23	.17	-.48	-.35	.11
13	Sodium (urine), 24 hour.....	7.85	4.58	.36	.18	-.88	-.47	1.00	-.30	-.21
14	Potassium (urine), 24 hour.....	3.20	1.17	.36	.09	-1.16	-.30	.86	-.13	.26
15	Serum phosphorus.....	3.55	.99	-.79	.44	.51	-.06	-.66	.07	.93
	Number of subjects.....	76	16	17	11	10	9	7	6

¹ Raw score units. Individual standard scores were based on these values.
² Reversed (200 minus actual time in centisecond.)

TABLE 3.—Summary of Classification of 7 Schizophrenic Types

	Number of subjects	Percent of 80	Percent of 76
Classified:			
Type 1, paranoid schizophrenic ¹	16	20	21
Type 4, paranoid schizophrenic ¹	10	12	13
Type 5, paranoid schizophrenic ¹	9	11	12
Total paranoid schizophrenic.....	35	44	46
Type 2, nonparanoid schizophrenic ¹	17	21	22
Type 3, nonparanoid schizophrenic ¹	11	14	15
Type 6, nonparanoid schizophrenic ¹	7	9	9
Type 7, nonparanoid schizophrenic ¹	6	8	8
Total nonparanoid schizophrenic.....	41	51	54
Total classified.....	76	95	100
Unclassified: Total unclassified.....	4	5
Total.....	80	100	100

¹ The types were called paranoid and nonparanoid on the basis of the diagnostic majority.

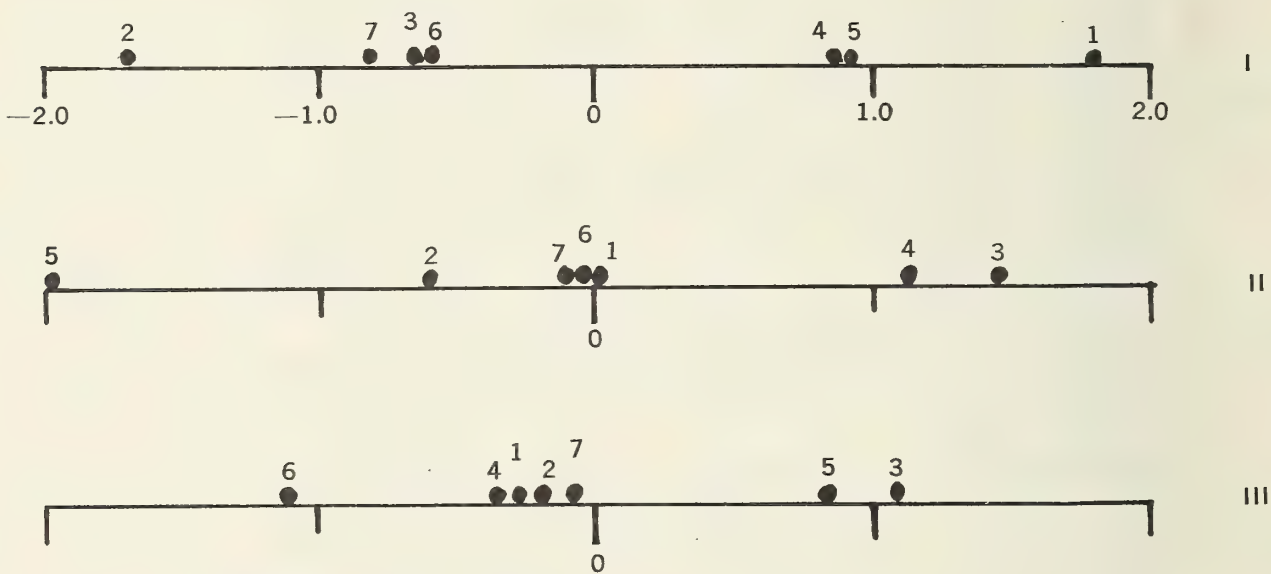


FIGURE 1.—Group means in one-dimensional discriminant spaces.

The first dimension (figure 1) which clearly separates the paranoid types (1, 4, and 5) from the nonparanoid (2, 3, 6, and 7) is defined by the performance on the facial expressions, graded analogies, and dotting speed tests, as well as by the parasympathetic index and the serum phosphorus level. This dimension resembles the extrovert-introvert continuum: outward oriented and active with low parasympathetic index and low serum phosphorus versus socially withdrawn and inactive with high parasympathetic index and high serum phosphorus. The second dimension (figure 1) is mainly defined by graded

analogies, level of thinking on the Kahn Test of Symbol Arrangement KTSA), autonomic reactivity, urine and sodium excretion, and the creatinine blood/urine ratio. It separates type 3 from the other nonparanoid types and the paranoid types, especially 4 and 5, from each other. All types, except 6 and 7, are in fact separated from each other on the first two dimensions (figure 2). Types 6 and 7 break apart on the third dimension (figure 1) which is defined by graded analogies, dotting speed, autonomic reactivity, the parasympathetic index, the creatinine blood/urine ratio, and the Humoller rate. Peculiar to this dimen-

TABLE 4.1.—Multiple Discriminant Analysis of 7 Schizophrenic Types

Function number	Normalized discriminant function coefficients for variable numbers															Percent of trace	Root	p<
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15			
1.....	0.50	0.34	0.16	0.48	-0.10	0.02	-0.12	-0.12	0.12	-0.28	-0.02	-0.12	-0.05	0.00	-0.47	59.5	4.806	0.001
2.....	-.01	.30	.37	-.10	.10	.14	-.28	-.32	-.26	-.13	-.41	.28	-.41	.21	.11	20.6	1.660	.001
3.....	.04	-.50	.11	.46	-.04	-.05	.14	.27	-.29	-.25	.21	.42	-.05	-.23	.02	10.3	.830	.001
4.....	.07	.31	-.42	.13	.26	-.45	-.21	.03	-.14	-.14	-.37	-.04	.02	.34	.30	5.1	.412	.025
5.....	-.46	.59	.20	-.20	-.39	-.11	-.03	.02	-.23	-.12	-.06	.04	.21	.17	-.24	2.8	.229
6.....	-.37	-.22	.21	.24	-.03	-.23	-.26	-.25	.19	-.51	-.39	-.20	.18	.08	.05	1.7	.140
Total trace.....																100.0	8.077	.001

Wilks' lambda for group discrimination=0.0179.
Which gives an $F=3.669$ on 90 and 316 degrees of freedom.
Total explained variance=54 percent.

TABLE 4.2.—Centroids in 3-Dimensional Discriminant Space

Dimension	Schizophrenic type						
	1	2	3	4	5	6	7
I.....	1.79	-1.69	-0.67	0.90	0.90	-0.66	-0.83
II.....	-.01	-.61	1.46	1.14	-1.98	.06	.09
III.....	-.27	-.21	1.08	-.33	.86	-1.11	-.12

(See also figs 1, 2, and 3.)

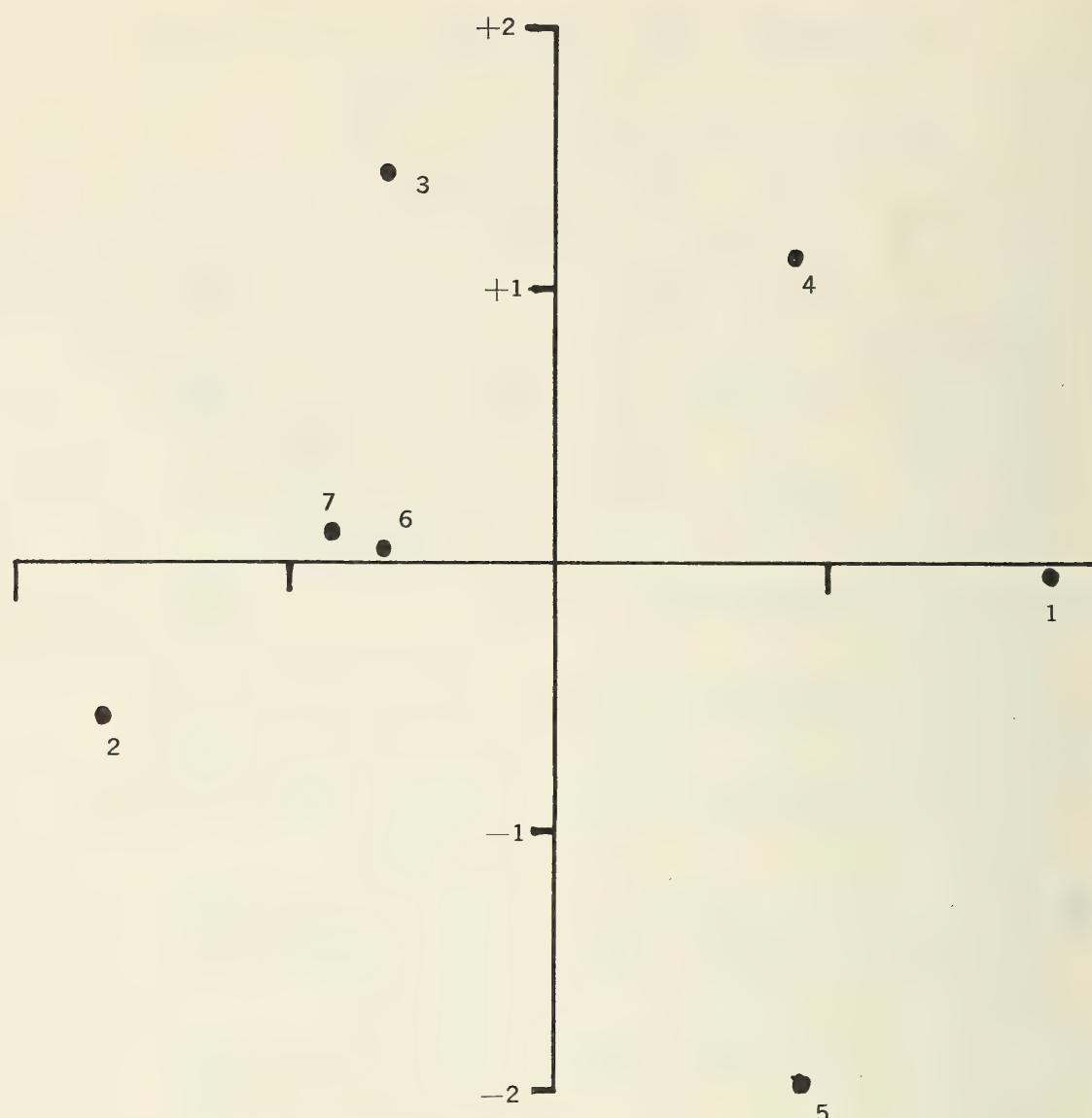


FIGURE 2.—Group centroids in two-dimensional discriminant space.

sion is, e.g., the combination of poor reasoning ability (graded analogies) with relatively fast psychomotor performance, and high heart rate increase with low blood pressure increase on epinephrine. Figure 3 demonstrates the group discrimination in three-dimensional space.

The next step in the analysis involved the computation of the best linear combination of the 15 variables for each of the 7 types (table 5). The seven linear discriminant scores equations could now be used as classification hypotheses for new individuals. These equations could also be applied back to the original sample to check for misclassifications as determined by the discriminant scores.

All seven discriminant scores were computed for each subject in the sample and the centroids for the types were obtained. As can be seen in table 6, the highest mean values, by a considerable margin, fall on

the diagonal of the distribution, indicating that each function is clearly associated with its type.

Table 7 shows how the subjects in each type were distributed by highest discriminant score. "Misclassification" (total, 13 percent) occurred, in all but one case, so that the mixup was either within the paranoid or within the nonparanoid types.

The summary descriptions of the seven schizophrenic types given below are based—in addition to the discriminant analysis—on univariate analyses of variance of all 388 project variables and on comparison of group means with grand means.

The differentiation of paranoid and nonparanoid types proved clearest on psychological test variables. In general, the performance of all paranoid types is above the mean and the performance of all nonparanoid types below the mean of the total schizophrenic sample. A ranking of the

TABLE 5.—Discriminant Score Weights for 7 Schizophrenic Types on 15 Variables

Variable No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	a
Type:																
1	0.357	0.270	0.148	0.339	-0.104	-0.008	-0.149	-0.151	0.098	-0.242	-0.021	-0.111	0.026	-0.022	-0.371	.703
2	-.452	-.190	-.112	-.383	-.004	-.006	.157	.135	-.041	.233	.047	.014	.064	.078	.291	.690
3	-.125	-.190	.139	-.023	-.018	.059	-.002	-.004	-.152	-.055	-.079	.205	-.140	-.124	.157	.477
4230	.333	.064	.081	.017	.021	-.116	-.129	-.044	-.061	-.092	.007	-.120	-.036	-.148	.374
5340	-.180	-.120	.339	-.063	-.029	.131	.213	.062	-.038	.273	-.038	.171	-.004	-.211	.775
6	-.099	-.091	-.053	-.258	.163	.104	.024	-.069	.140	.212	-.010	-.057	-.036	.003	.135	.406
7	-.220	-.015	-.196	-.118	.196	-.182	-.074	.003	-.048	-.002	-.177	-.008	-.008	.127	.284	.458

Discriminant score for subject (*i*) from function of subtype (*k*):
 $D_{ik} = w_{1i}z_1 + w_{2i}z_2 + \dots + w_{15i}z_{15} - a_k$,
where w_1, w_2, \dots, w_{15} are the weights for subtype (*k*) and
 z_1, z_2, \dots, z_{15} are the values in standard score units for subject (*i*) on the 15 variables.

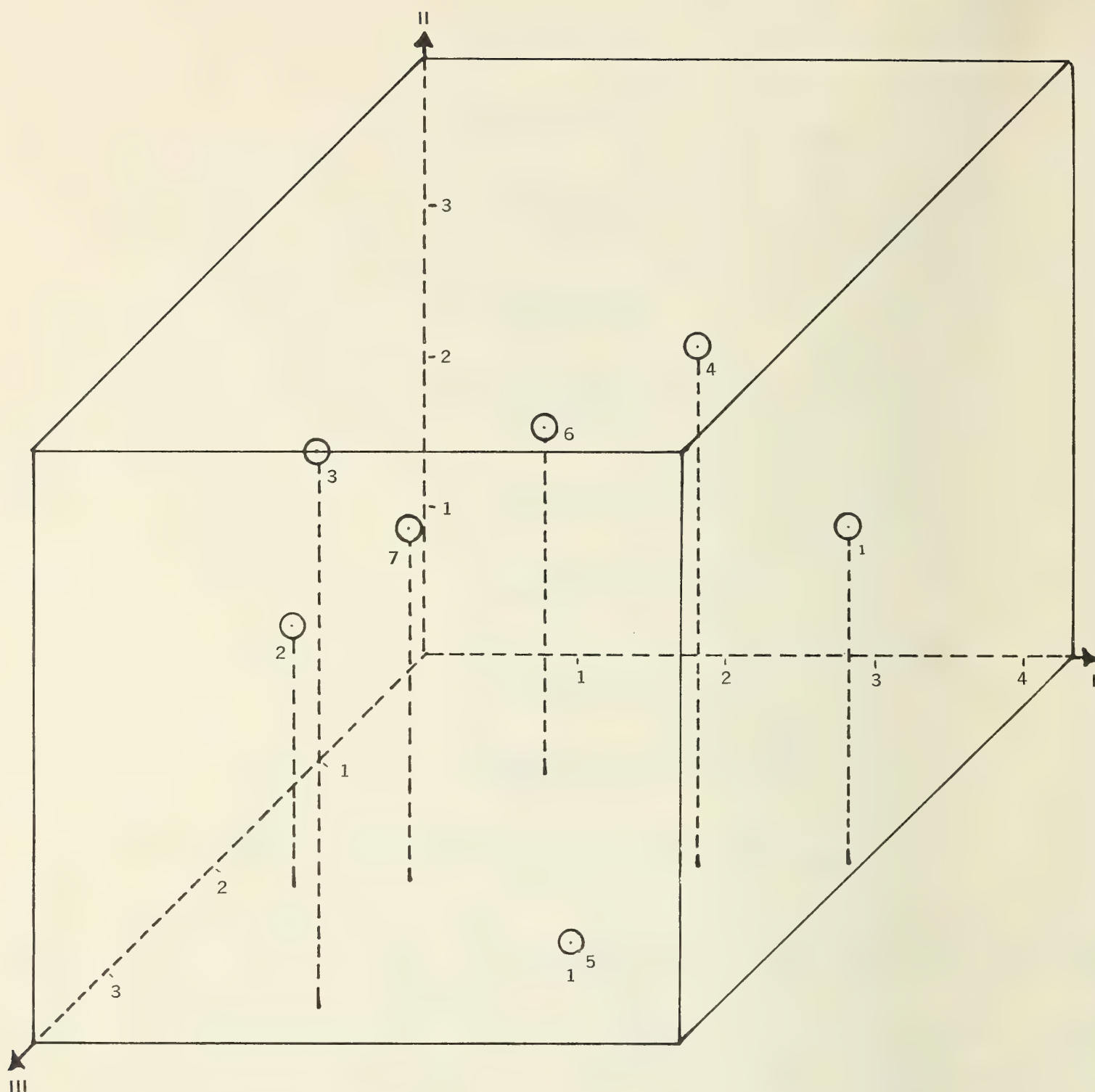


FIGURE 3.—Group centroids in three-dimensional discriminant space (2.5 added to each score).

means on 70 psychological test variables revealed the following sequence for the seven types in regard to their general level of functioning:

	Poor performance				Good performance		
Type number	2	3	6	7	5	1	4
Diagnostic majority ¹ .	NP	NP	NP	NP	P	P	P

¹ NP=Nonparanoid, P=paranoid schizophrenic.

Psychiatric ratings consistently describe the paranoid types as manic, agitated, belligerent, less withdrawn, and conceptually

better organized in contrast to the non-paranoid types who are depressed, nonagitated, submissive, withdrawn, and conceptually more disorganized.

Patterns of childhood adjustment which most clearly differentiate the two sets indicate that the paranoid types were more alert in childhood, went farther in school, did better scholastically, and had better social adjustment.

Paranoid types

Type 1

N=16, mean age=35, mean hospital stay=9 months

TABLE 6.—Mean Discriminant Score ¹ for 7 Schizophrenic Types on all 7 Discriminant Scores Functions

Schizophrenic type		Discriminant score function used						
		No. 1	No. 2	No. 3	No. 4	No. 5	No. 6	No. 7
1. Paranoid	Mean	0.704	−1.951	−1.006	0.285	−0.184	−0.846	−1.139
	S.D.	.565						
2. Non paranoid	Mean	−1.963	.690	−.252	−1.148	−1.257	.102	.064
	S.D.		.523					
3. Non paranoid	Mean	−1.232	−.464	.477	−.421	−1.401	−.471	−.340
	S.D.			.539				
4. Paranoid	Mean	−.046	−1.462	−.525	.375	−.986	−.586	−.661
	S.D.				.334			
5. Paranoid	Mean	−.110	−1.174	−1.102	−.584	.774	−.933	−.881
	S.D.					.543		
6. Non paranoid	Mean	−1.144	−.236	−.541	−.554	−1.304	.405	−.303
	S.D.						.239	
7. Non paranoid	Mean	−1.384	−.169	−.360	−.557	−1.197	−.252	.458
	S.D.							.265

¹ The means were calculated from the individual discriminant scores for the subjects in each group. The off-diagonal values are given for comparison. Standard deviations were calculated only for the diagonal.

TABLE 7.—Table Showing How the 76 Schizophrenic Subjects Were Distributed According to Their Highest Discriminant Score, by Type and Function

Schizophrenic type		Discriminant function no.							N Total	Pro. right
		1	2	3	4	5	6	7		
1. Paranoid		12			2	2			16	0.75
2. Nonparanoid			15				1	1	17	.88
3. Nonparanoid				10	1				11	.91
4. Paranoid		2			8				10	.80
5. Paranoid		1				8			9	.89
6. Nonparanoid							7		7	1.00
7. Nonparanoid								6	6	1.00
Total									76	.87

A total of 10 subjects were misclassified according to their discriminant score (only 1 was misclassified as to paranoid or nonparanoid type in general).

Psychology.—The patients in type 1 are among the most efficient of all patients (schizophrenics and nonschizophrenics) in regard to overall psychological functioning. Compared with the other schizophrenic types they score highest ⁸ on 9 and second highest on 30 out of 70 test variables. Intellectually well within average normal limits (the mean Wechsler-Bellevue full IQ=107.0 ⁹), their only noticeable deficiency is a slight slowing of reaction time (mean

RT=23.6 centiseconds as compared with normal mean RT=18.5).

Psychiatry.—Like the other paranoid types these patients are manic, agitated, belligerent, and little withdrawn. They differ in being the most compliant and least anxious of all schizophrenics. They also manifest a high level of perceptual distortion and are fairly active.

Social history.—These patients come from higher social class with relatively little reported pathology in their family backgrounds. Although they had childhood health problems, they were alert, did well scholastically, went farthest in school, and were best adjusted in their social relations

⁸ In using superlative and comparative forms to describe characteristics of a particular type, the reference is to the other schizophrenic types unless otherwise indicated.

⁹ With the exception of the digit symbol, the W-B subtests or IQ's were not used in sorting patients into types.

with peers in childhood and adolescence. They were, however, the most brooding as children and had feelings of personal failure and inadequacy in late adolescence. Both parents were adequate, the father being the more dominant figure in the family.

Anthropometry.—Tallest of all schizophrenics, with relatively long legs, they also have the highest Androgyny score (89.5).¹⁰

Physiology.—On autonomic nervous system tests these patients have the smallest resting blood pressure standard deviation (autonomically most stable) and the lowest parasympathetic index. Heart rate and blood pressure change very little in response to piperoxan and yohimbine, whereas there is a large blood pressure increase in response to epinephrine and norepinephrine. On mecholyl the blood pressure fall is minimal, smallest of all types.

Biochemistry.—Biochemical tests show that this type has the lowest serum phosphorus level, high creatinine excretion, and the highest norepinephrine excretion, especially in relation to 5-HIAA.

Type 4

N=10, mean age=31, mean hospital stay=17 months

Psychology.—Type 4 scores highest on more test variables (27) than any other type. Their performance is good especially on intellectual tests; e.g., W-B full IQ=109.4. There is, however, evidence of some psychomotor retardation extending from slight slowing of reaction time (21.3 centi-seconds) to a more noticeable slowing of continuous motor performance (e.g., dotting speed). Although their performance on stressful tasks, such as the Stroop Color-Word Test, is better than that of any other type, they experience difficulty when the test procedure departs from the ordinary: slow blindfolded performance on the

Seguin Formboard, inaccurate on the Mirror Drawing Test, part II, where instructions are reversed from those in part I, and relatively slower and inaccurate on the stressful second part of the Stroop Color-Word Test.

Psychiatry.—In addition to being manic, agitated, and little withdrawn like the other paranoid types, these patients are most belligerent and conceptually most intact of all schizophrenics. They are also fairly inactive and have grandiose delusions.

Social history.—These patients come from higher social class and from relatively stable and cohesive families. They had good health in childhood, were alert, aggressive, and did well in school. While their fathers were adequate, their mothers were inadequate: un insightful, submissive, and interested in very little outside the home. The patients, although dependent on their mothers, felt that their mothers considered them unacceptable. The quality of this relationship bears resemblance to Bateson's "double bind" concept.

Anthropometry.—The Androgyny score is second highest (88.9).

Physiology.—The psycho-galvanic reflex (PGR)¹¹ baseline is the highest of all. Heart rate and blood pressure change very little in response to piperoxan and yohimbine (lowest of all). They have the largest heart rate increase and the second largest blood pressure increase on atropine.

Biochemistry.—This type has the highest BSP percent retention and highest total catecholamine excretion.

Type 5

N=9, mean age=34, mean hospital stay=16 months

Psychology.—The type 5 patients, ranking third on overall psychological test performance, score highest on 16 and second highest on 8 test variables. They are especially good on some psychomotor tests (dotting, tapping, and Bourdon Dot Can-

¹⁰ The Androgyny score, a body index of masculinity, is calculated as three times the shoulder width (cm.) minus the pelvic width (bone measures). The normal male average is 100.

¹¹ Refers to the production by the body of a weak current on the skin surface.

cellation) on which their performance is faster than that of any other patients including the nonschizophrenics. Intellectually within average normal range (W-B full IQ=101.4), they do better on nonverbal performance tests (W-B performance IQ=109.4) than on verbal tests (W-B verbal IQ=93.8), a discrepancy not present in types 1 and 4. They are form reactive in their response to Thurstone's Color-Form Movie, perform well on such tests as Mirror Drawing, Seguin Formboard (blindfolded), and kinesthetic aftereffect, and less well on verbal reasoning and concept formation tests, such as graded analogies, Kahn Test of Symbol Arrangement, picture group naming, and word classification.

Psychiatry.—According to psychiatric and ward ratings these patients are the most manic, most agitated, least withdrawn, most hyperactive, most euphoric, most self-depreciative, and very anxious.

Social history.—As in type 1, the patients in type 5 had considerable childhood health problems. They were less alert than the other paranoid types and socially less well adjusted in childhood and adolescence. Worried and brooding, they were anxious to avoid conflicts with peers and authority figures. Their self-evaluation was unrealistically high in late adolescence. Both of their parents had severe personality problems and were resentful of each other. Especially the fathers were inadequate and insecure. Although dependent on both parents, these patients were more dependent on their fathers than any other patient type.

Anthropometry.—Anthropometric measurements indicate a generally small and lean body build.

Physiology.—The galvanic skin response (GSR)¹² baseline is highest and the psychogalvanic reflex (PGR) baseline lowest of all patient types. They show the highest increase in PGR, respiration rate, heart

rate, and blood pressure, and a large drop in GSR in response to shock. They have a large increase in heart rate and blood pressure on piperoxan and yohimbine, highest heart rate increase on epinephrine, and least heart rate decrease on norepinephrine.

Biochemistry.—These patients have the highest urine volume (2.4 liters/24 hours) and the highest per volume excretion of sodium, potassium, and creatinine. The total catecholamine excretion is lowest.

Nonparanoid types

Type 2

N=17, mean age=37, mean hospital stay=53 months

Psychology.—The type 2 patients are the most inefficient of all: their functioning on psychological tests is even below that of the brain-damaged control patients. Of borderline intelligence (W-B full IQ=74.8), they score lowest on 54 test variables. In fact, they do not reach an adequate level on any test. Their reaction time is extremely slow (50.3 centiseconds).

Psychiatry.—All signs of overt psychosis are present: Highest resistiveness; most perceptual distortion; most motor disturbance; most grandiose expansiveness; and highest mental and conceptual disorganization.

Social history.—These patients come from low social class with enormous amounts of pathology in their family backgrounds. They were the most unalert of all patients in early childhood, had low education level, poor scholastic performance, and poor social adjustment throughout childhood and adolescence. They had low self-evaluation, poor frustration tolerance, and were aggressive with peers and extra-family authority figures. The parental home was extremely disorganized and filled with hostility. Both parents were very inadequate. The fathers were least insightful and most domineering, drank most heavily, and punished most violently. The mothers were also least insightful and away from home more than the mothers of any other type.

¹² Refers to the apparent resistance of the skin to the passage of a weak external electric current.

Anthropometry.—The body build is asthenic: narrow shoulders, narrow chest, relatively long arms, and low Androgyny score (79.9).

Physiology.—These patients have the largest resting blood pressure standard deviation (high autonomic lability). In response to mecholyl they have the largest heart rate increase of all patient types and the second largest fall in blood pressure. Their response to other experimental drugs reveals no specific trends.

Biochemistry.—The 24-hour total epinephrine excretion is highest of all patient types.

Type 3

N=11, mean age=37, mean hospital stay=66 months

Psychology.—Intellectually these patients are within dull normal limits (W-B full IQ=88.5) overlapping with the brain-damaged control subjects. They score lowest on 2 and second lowest on 33 out of 70 test variables. Their reaction time is extremely slow ((50 centiseconds). In their response to Thurstone's Color-Form Movie they are color reactive.

Psychiatry.—Psychiatric ratings describe these patients as conceptually disorganized, self-depreciative, most depressed, and most withdrawn.

Social history.—These patients come from a social class higher than that of any other patient type. They were relatively unalert in early childhood and had about average social adjustment which deteriorated toward adolescence. They were then withdrawn, conforming, and nonaggressive, had extremely low self-evaluation but fairly good frustration tolerance. Their fathers were successful in and outside the home, while their mothers were relatively disorganized, inadequate, and uninsightful. This resulted in some conflict in the family.

Anthropometry.—No specific trends are present.

Physiology.—The heart rate increase is the second highest and the blood pressure

increase the lowest of all patient types on piperoxan and ephinephrine.

Biochemistry.—These patients have the lowest urine volume (less than 1 liter/24 hours) and the lowest per volume excretion of sodium, potassium, creatinine, and nor-epinephrine. They have the highest creatinine blood/urine ratio and the highest Humoller (serum oxidase) activity (O.D. at 6 minutes).

Type 6

N=7, mean age=39, mean hospital stay=101 months

Psychology.—The functioning of this group of patients on psychological tests is generally below the schizophrenic average, especially on intellectual tests (W-B full IQ=89.7, within dull normal range). On psychomotor tests they score higher: their reaction time (24.0) is close to that of the better functioning paranoid types (1, 4, and 5) and their performance on hand-eye coordination tests (mirror drawing, tweezer dexterity) is particularly good.

Psychiatry.—These patients are on the depressed side, not anxious, not angry, compliant, most submissive, least motor disturbed, and perceptually and conceptually relatively intact.

Social history.—The social background of this type resembles very much that of type 2: low social class, low education level, poor school performance, poor social adjustment in childhood and adolescence, very low frustration tolerance, and very poor occupational adjustment. Their homelife was apathetic, and both parents were inadequate in their parental roles: mothers, although more adequate than fathers, were uninsightful and submissive, and fathers were heavy drinkers and irresponsible as heads of their households.

Anthropometry.—No outstanding features.

Physiology.—The patients in this type have the highest parasympathetic index, and the highest skin temperature baseline, mean=31.6° C. (total research sample

30.2° C.). In response to epinephrine they have the smallest heart rate change (it actually drops) and the largest blood pressure increase. On norepinephrine they have the largest drop in heart rate and the largest increase in blood pressure. They have the second lowest increase in heart rate and the lowest increase in blood pressure in response to atropine.

Biochemistry.—The glucose tolerance values (fasting level, absorption rate, and maximum utilization rate) are highest for this type. They have the lowest norepinephrine excretion in relation to 5-HIAA (24-hour ratio and milliliter ratio).

Type 7

N=6, mean age=33, mean hospital stay=30 months

Psychology.—Of the nonparanoid patients type 7 is psychologically the best functioning, scoring frequently at or above the schizophrenic average. Their intelligence is within average normal limits (W-B full IQ=100.2) but they have difficulties in symbolic thinking and reasoning as indicated by the Kahn Test of Symbol Arrangement. They perform rather well on perceptual and psychomotor tests. They have fast and accurate form perception (identical forms and changing pictures), good closure ability (Mooney Closure Faces Test), accurate kinesthetic aftereffect, and good orientation to true vertical (rod and frame). Although slightly slower than normal, their reaction time is fastest of all types (20.2). Their performance tended to improve when they were urged to go faster (dotting, tapping) or when a stress factor was introduced (Bourdon Dot Cancellation, Stroop Color-Word Test).

Psychiatry.—Psychiatric and ward ratings describe these patients as submissive, angry (but not belligerent), underactive (least active of all), least agitated, mentally relatively well organized and most anxious.

Social history.—These patients come from stable families of medium low social class. They had good childhood health,

good social adjustment in childhood and adolescence, very good frustration tolerance, and a relatively good job record. In adolescence they were fearful of women. The patients remained close to and dependent on their mothers who were gentle and supportive with better than usual insight. The fathers, who drank excessively, were hostile, dependent, and weak.

Anthropometry.—No trends.

Physiology.—The patients in this type have the lowest baseline skin temperature (mean=26.5° C.), highest heart rate and blood pressure increase on yohimbine, second smallest drop in heart rate and smallest increase in blood pressure on norepinephrine, and second largest heart rate increase and largest fall in blood pressure in response to mecholyl.

Biochemistry.—The serum phosphorus level is highest and the glucose tolerance maximum utilization rate is lowest of all types. The per volume epinephrine excretion is highest of all.

The seven schizophrenic types were found, as already has been pointed out, on the basis of the individual patients' scores on a small number of objective psychological, physiological, and biochemical tests and measurements. Nevertheless, these types have meaningful, distinct characteristics also in terms of subjective evaluations and ratings. For example, psychiatric and ward behavior ratings might suggest the following labels:

Type 1. Compliant paranoid schizophrenic.

Type 2. Overtly psychotic nonparanoid schizophrenic.

Type 3. Depressed nonparanoid schizophrenic.

Type 4. Belligerent paranoid schizophrenic.

Type 5. Anxious and agitated paranoid schizophrenic.

Type 6. Submissive nonparanoid schizophrenic (burnt-out residual).

Type 7. Anxious and underactive non-paranoid schizophrenic.

Similarly, labels could be attached to the types in terms of social history information. But, it is not the pattern of characteristics in one particular area of study that makes a type; it is the combined pattern across areas. Many individuals may fit the description of a type in terms of characteristics in a particular area, but may not fit the description of this type in another area.

It is interesting to note that age was not a contributing factor in the discrimination of the seven types. The mean ages ranged from 31 to 39. The paranoid types had shorter mean stay in hospital than the others. Yet, the fact that the paranoid types were clinically better integrated and psychologically better functioning, does not necessarily depend on the shorter duration of hospitalization. It may rather depend on the nature of the illness, as indicated by the fact that the two least well functioning types (2 and 3) did not have the longest hospitalization. The role of age and length of hospitalization was probably more evident in the *Q*-factor types because of the mixture of schizophrenics and nonschizophrenics. For example, in the positive group of *Q*-type 8 age-linked liver function deficiency is present, but the group consists of older nonschizophrenics with alcoholism. Length of stay in hospital seems to follow the schizophrenic—nonschizophrenic and the paranoid—nonparanoid schizophrenic divisions. The nonschizophrenics, in general, had a shorter duration of hospitalization and divided, for the most part, into younger personality disorders or older alcoholics.

The patients' race and ethnic origin were not correlated with either the *Q*-factor types or the seven schizophrenic types. The proportions of Negroes, e.g., in the seven types were 0.25, 0.40, 0.20, 0.30, 0.11, 0.30, 0.33, respectively.

Compared with the *Q*-factor types (total patient sample), number 1, 2, and 5 of the

seven schizophrenic types correspond fairly well to *Q*-1 positive, *Q*-1 negative, and *Q*-6 positive in the order mentioned. The well-known paranoid-nonparanoid division, also evident in the *Q* analysis, gets even stronger support from the seven types. This division, as well as the further break of these major groups into subgroups, was a result of statistical sorting of objective data, not a consequence of clinical diagnosis. The types were called paranoid or nonparanoid on the basis of the diagnostic majority. An individual does not necessarily have to carry a clinical diagnosis of paranoid schizophrenia to be classified into such a type.

With reference to the classic Kraepelinian and Bleulerian types: (a) Paranoid schizophrenia divided into three distinctly separate categories (types 1, 4, and 5); (b) no types could be identified as clearly catatonic or hebephrenic, probably because these types were missing from the sample (only one patient was diagnosed as catatonic and one as hebephrenic, both of them appearing in type 3); and (c) a fairly good picture of simple schizophrenia could be seen in type 7 (despite the fact that only one of the four patients diagnosed as simple schizophrenic fell into type 7, the rest falling into type 2).

In addition to separating the psychologically generally well-functioning paranoid schizophrenics from the generally poorly functioning nonparanoid schizophrenics the seven types also support or reflect more specific trends reported in the literature; e.g.,

1. Wentworth's:

Group 1: Uniformly poor intellectual test performance (types 2, 3, and 6);

Group 2: Variable intellectual test performance (types 5 and 7);

Group 3: Uniformly efficient intellectual test performance (types 1 and 4).

2. Beck's:
 - S-1: Intellectual disruption, advanced stage of disease, a very sick pattern, sickest of all (type 2);
 - S-2: Affect-dominated, ego control lacking, intellectual disruption, also advanced stage of disease (type 3);
 - S-3: Orderly intellectual functioning, healthy appearing exterior, defensive (types 1 and 4);
 - SR-2: Orderly intellectual functioning, autistic and withdrawn but striving to maintain balance (type 7).
3. Bateson's "double bind" patient-mother relationship (type 4).
4. Rodnick and Garmezy's:
 - (a) good premorbid adjustment type (especially types 1 and 4 of the better adjusted 1, 4, 5, and 7);
 - (b) poor premorbid adjustment type (especially type 6 of the less well adjusted 2, 3, and 6).
5. Lindemann's:
 - (a) active schizophrenic showing marked cardiovascular changes on epinephrine (type 5);
 - (b) apathetic schizophrenic showing little response to epinephrine (type 6).
6. Holmberg's anxious yohimbine responder (types 5 and 7).
7. Funkenstein's:
 - (a) mecholyl nonresponder with high norepinephrine excretion (type 1);
 - (b) mecholyl responder with high epinephrine excretion (types 2 and 7).
8. Deviant carbohydrate metabolism—possibly corresponding to Meduna's oneirophrenia (types 6 and 7).

Concluding Remarks

As shown by our study, there is sufficient connection between behavioral and biological functioning to support typologies of people on the combination of the two. We

have presented two typologies of which the second one, the seven schizophrenic types, is sharper and better defined. From their description they seem to reflect current trends and views as regards schizophrenia, and mathematically they appear to be distinct types separated by more or less sharp boundaries in test space. The probability levels attached to the power of discrimination are, naturally, unreliable due to such factors as: (1) The sampling method, intended to be random, becoming biased in favor of numerous selection criteria and other necessary considerations; (2) the sample size, small to begin with, resulted in types containing too few members to yield stable estimates of centroids and dispersions. The sampling method also restricted the heterogeneity of the sample and limited it to start with into a, so to say, type of patients who fulfill certain requirements. The time element, although most certainly in some way significantly affecting the shape of the test profiles, was not considered.

The obtained typologies show clearly that characteristics commonly attributed to schizophrenics as a whole may apply only to certain subgroups of schizophrenia.

In this study our intention was to find sufficiently distinct types of schizophrenia on which to build in future research, including etiology, pathology, treatment, and prognosis. It was, indeed, our intention to conduct such validation studies later but they did not materialize. It is hoped that our findings will be of service to future investigators.

Acknowledgment

The summary descriptions of the types are based, to some extent, on memoranda written by project staff members. Extensive further aid, however, in the examination and interpretation of the actual results as well as in the present writeup of the descriptions was received from the project psychologist, Eira I. Mattsson.

REFERENCES

- (1) BURT, C. L. "Correlations Between Persons." *Brit. J. Psychol.*, 1937, 28, 59-96.
- (2) CATTELL, R. B. "Factor Analysis." New York: Harper & Brothers, 1952.
- (3) COOLEY, W. W., and LOHNES, P. R. "Multivariate Procedures for the Behavioral Sciences." New York: Wiley & Sons, 1962.
- (4) GERARD, R. W., and MATTSSON, N. B. "The Classifications of Schizophrenia." In J. A. Jacquez (ed.) "The Diagnostic Process." Ann Arbor, Mich.: The University of Michigan, 1963.
- (5) HARMAN, H. H. "Modern Factor Analysis." Chicago: The University of Chicago Press, 1960.
- (6) HOLZINGER, K. J., and HARMAN, H. H. "Factor Analysis." Chicago: University of Chicago Press, 1941, p. 28.
- (7) KAISER, H. F. "The Varimax Criterion for Analytic Rotation in Factor Analysis." *Psychometrika*, 1958, 23, 187-200.
- (8) KENDALL, M. G. "Rank Correlation Methods." London: Charles Griffin & Co., 1958.
- (9) KENDALL, M. G. "Discrimination and Classification." Paper read at the International Symposium on Multivariate Analysis, Dayton, Ohio, June 16, 1965.
- (10) KOSSACK, C. F. "Statistical Classification Techniques." *IBM Systems Journals*, 1963, 2, 136-151.
- (11) NEUHAUS, J., and WRIGLEY, C. "The Quartimax Method: An Analytical Approach to Orthogonal Simple Structure." *Brit. J. Stat. Psych.*, 1954, 7, 81-91.
- (12) RAO, C. R. "Advanced Statistical Methods in Biometric Research." New York: Wiley & Sons, 1952.
- (13) RICE, C. E., and MATTSSON, N. B. "Types Based on Repeated Measures." In M. Lorr (editor) "Explorations in Typing Psychotics." New York: Pergamon Press, 1966.
- (14) ROGERS, D. J., and TANIMOTO, T. T. "A Computer Program for Classifying Plants." *Science*, 1960, 132, 1115-1123.
- (15) SAUNDERS, D. R. "Some Varieties of Schizophrenia." In "Symposium: Explorations in Typology With Special Reference to Psychotics." New York: Human Ecology Fund, Inc., 1964.
- (16) STEPHENSON, W. "The Inverted Factor Technique." *Brit. J. Psychol.*, 1936, 26, 344-361.
- (17) THOMSON, G. "The Factorial Analysis of Human Ability." University of London Press, 1950.
- (18) TUCKER, L. R. "A Method for Synthesis of Factor Analysis Studies." "Personnel Research Section Report, No. 984." Washington, D.C.: Department of the Army, 1951, p. 120.
- (19) ZUBIN, J., and FLEISS, J. L. "Taxonomy in the Mental Disorders—A Historical Perspective." In "Symposium: Explorations in Typology With Spe-

cial Reference to Psychotics." New York: Human Ecology Fund, Inc., 1964.

Supplementary Comments by Dr. Gerard as Presented at the Conference

The paper before you presents the analytical procedures and a summary of the crucial outcomes, especially Mattsson's important methodological contribution of "successive factorial screening." You have been referred in it to earlier publications which give in somewhat more detail the framework from which these results emerged; I urge attention especially to the paper in Behavioral Science. I regret that the full data, and the full presentation of the instruments with which these were obtained, are not yet available. A monograph is reasonably nearly done, but is in abeyance at present. My role now will be to place our work in the broader framework of this meeting, and I shall do so with a series of rather categorical statements, since the arguments pro and con are now clearly before us. Oddly enough, it seems to me logical to present the material in reversed order.

First, a statement of fact; our concern in this work was entirely that of research and understanding, with the firm conviction that such work will in time lead to important uses.

Second, we affirm our commitment to the assumption that there exist meaningful disease entities, in the sense that there are patterns of phenomena running together, syndromes which sooner or later will have prognostic value, etiological value, and the like. The proof of such outcomes is often very delayed. In biology the concept of species and classification, as worked out by Linnaeus largely, really did not get the kind of firm substantiation that it now has until genes and amino acid sequences and proteins and immunochemistry came along

and showed that, indeed, in terms of the amino acid shifts in proteins, dogs and foxes are more related than dogs and bears, and these more than both dogs and bears versus the cat family.

Given this assumption, progress in this area, in all of these areas, is truly Operation Bootstrap. One starts by observing the world, dividing it as best one can into some set of classes, and seeking criteria which help discriminate these. This leads to improving the classes, improving the criteria, and iteratively doing better and better until it is possible to make an hypothesis or a model with some real power. From then on the advance is usually explosive and very much more satisfactory.

The start, however, is with a naturalist in biology. In the mental area the clinician is the naturalist—he observes the range of phenomena, asks the questions, and is the artist who has the hunches. In terms of his hunches about the way the universe of mental illness might divide, I have the greatest respect for him. His hunches, however, as those of any other observer in the world, must be subject to verification or deverification. The questions he raises must be answered, and the chosen ways of doing so are sound or unsound. At this stage one cannot trust the clinician, or the naturalist, too greatly; this is where the analytic scientist must slowly and painfully work out the correct answers.

Our work was based on the assumption that clinicians did have some basis for separating mental patients into schizophrenics and nonschizophrenics. We would, therefore, examine a large array of criteria, potentially able to distinguish these classes, in all fields—from psychiatric and social history through ward observations to psychological, anthropometric, biochemical, and physiological measurements—the entire spectrum. We made the further assumption that those criteria which served more or less to discriminate schizophrenics from nonschizophrenics would also prove useful in separating schizophrenic 1 from

schizophrenic 2, if there were indeed such subgroups. This was our procedure.

The outcome you have before you. Having made a great matrix of criteria, we did our final analysis with a very limited number of them, all chosen from the fields of objective measures—psychology, biochemistry, physiology. These divided our schizophrenic population into seven groups, which may or may not have had any meaning at all, even though the groupings obtained under somewhat different analyses were reasonably consistent. The fact, however, that when we looked at the data from psychiatry and from social history and ward behavior of the patients already assigned to the seven groups, these also fell into clinically meaningful clusters or syndromes or descriptions, is an internal kind of validation of considerable worth. Indeed, some participating psychiatrists, when they first saw the results, were extremely excited; one said, “You have reproduced significant psychiatric types.” Of course we had hoped to move on and further test our subgroups by longitudinal studies of natural changes in time, by studies of responses to drugs and to other new variables, and the like; but time ran out.

Now I shall back up a bit to what we did methodologically. When one does not have an intellectual model of sufficient vigor and reasonableness to permit rational experimentation, then one can only do what may properly be called “idiot research.” We did idiot research; we said, “Let’s look at the world and see what we can find.”

Now I submit that the computer is an extremely effective idiot. It does enable one to take a great mass of noisy data and from that dissect out the signals which are consistent and reproducible. I remind you, in the field of the EEG, for example, where the data are extremely noisy, where the interpretation of the wiggly, nonquantitative, analogical curves was at first purely by artistic clinical impression; that now

such data are analyzed and digitalized and successfully handled by computers. It is possible, even, with autocorrelation or by averaging large numbers of timed records, as an example, to demonstrate evoked potentials in the brain which nobody had previously been able to see in the raw recordings. I further suggest that this is exactly what the empirical computer approach is able to do with the noisy clinical data of such concern to Dr. Lehmann. The computer, used wisely, can return to the clinician better material than he gave it, by picking out certain consistent meaningful relationships, certain signals, from a great deal of confused noise.

I also must point out an interesting paradox. We all recognize that, in the progress of science, it is always possible to disprove an hypothesis by negative results, but it is not possible to prove something for all time, since new opposing facts may turn up to negate the interpretation. But in the case of identifying clusters or subtypes or groups, whatever word you prefer, the situation is rather reversed. If one does find groups that prove to be useful and which meet some of the criteria of consistency and of correlation with other variables, then such groups are there. They may not be the best ones, but they can be found. If one doesn't find useful groups it does not prove that they are not there but only that one has not yet been smart enough to succeed.

Consider now the data, themselves. Although we started with the assumptions that there are disease entities and subtypes of schizophrenia that one might find, we were like the man who said he didn't trust in God for, no matter how bright the sunshine, he carried an umbrella. We did include measurements on nonschizophrenics, and intend to publish the complete array of all observations on every patient in rather extensive tables. Whether our interpretation and patterning of them prove

ultimately really useful or less useful, the data themselves must remain the valuable base from which others can work.

The observations were made simultaneously, as nearly as we could, to avoid shifts between phases of a varying clinical state; and they were longitudinal to the extent that we did include the social history. (And I remind you that a tadpole is more like a frog, though they look very different, than a frog is like a toad, though they look very much alike, in terms of classification.)

The procedures used are sufficiently presented in the papers. I do emphasize that we took great care to control all the variables that we could think of, so that whatever quantitative findings emerged (and all subjective measurements were rated and, thereby, quantified) represented, as nearly as possible, groups of subjects separated by their illnesses and not by unrecognized environmental inputs. The two groups, schizophrenics and nonschizophrenics, were always run through the research ward together in each weekly subunit; and controls on the psychological environment, the drug environment, the temperature environment, the diet environment, etc., were as good as we could make them. We do have, as was emphasized earlier, a limited part of the total population of schizophrenics; only those in a given kind of hospital. If anything, we have thus lost some possible subtype, and we have certainly not created any. But only under hospital conditions could adequate environmental controls be established.

So, in summary, we deliberately used a very large matrix—not because patients are expensive and we want to get as much as possible out of them, but because the larger the matrix the more the potential outcome. We did get subtypes which seemed meaningful on respectable criteria, and I don't think that the route to get there, although pretty roundabout, was unnecessarily so.

Neurophysiological Response Strategies in the Classification of Mental Illness¹

Max Fink, M.D.²

In the absence of specific etiological theories of mental illness, physiological responses provide a basis for systems of classification. The selection of the physiological response measure is empiric, dependent on the experience and interests of the researcher and the types of available instrumentation. Because the conditions of testing and criteria of change are varied, the results appear inconclusive and few studies are comparable. That no technique has been satisfactory is attested by the lack of general application despite much effort. Perhaps for these reasons, most studies have been dedicated to a more pragmatic goal—that of identifying patients who are treatment-sensitive and selecting appropriate therapies.

But a theoretic model is implicit in the attempts at classification of the mentally ill by their physiological responses. The theoretic construct of an infectious etiology was the basis for the identification of neurosyphilis; and that of a metabolic defect in phenylketonuria. The clinical patterns (the phenotypic responses) are not wholly discriminatory, and the identification of each subgroup depends on a single criterion for the genotypic response. Implicit in studies of neurophysiological responsivity is the model that cerebral dysfunction or altered sensitivity is the basis of the expression of the functional psychiatric diseases, and that a specific response difference exists for each subgroup.

In these studies, the autonomic response to cholinergic and adrenergic drugs, and

the EEG responses to sedative and convulsant drugs and to sensory stimuli have received the most attention. The blood pressure response to methacholine and epinephrine—popularly described as the Funkenstein test—are representative of the autonomic studies (2). The EEG response to intravenous barbiturates as in the sedation-threshold or sleep-threshold studies of Shagass (3), the pentothal activation studies of Goldman (4), Itil (5), and Roth (6), and the chloralose studies of Monroe (7) are the better known of the EEG studies. Other measures that have received extensive study are the visual and somesthetic evoked cortical response (8), pupillary response (9), conditionability of eyeblink (10), visual critical flicker fusion (11) and the galvanic skin response (12).

To these response measures should be added a long list of stressors including sedative, stimulant, psychotropic, and hallucinogenic drugs (13), repetitive sensory stimuli (8), perceptual isolation (14), and sleep (15). This list and some representative citations are presented to indicate that the problem has received considerable attention, but for the sake of a focus, the review will summarize the data of some EEG-response studies in relation to the problem of classification of the mentally ill.

The resting electroencephalogram has been examined on the assumption that the laboratory situation itself represented a standardized stress and that its variability is representative of the subject's clinical diagnosis (16). Early studies indicated that certain psychotic subgroups have choppy rhythms, while recent determinations suggested that schizophrenic subjects have greater variability in amplitude measures than normal subjects. Schizophrenic subjects have been reported to have greater amounts of fast activity than depressives, but also greater amounts of theta-delta activity. Contrasting different regions of the scalp, observers have reported poor synchronization, poor regional organization, and low correlations in schizophrenic

¹ From the Department of Psychiatry, New York Medical College, New York, N.Y. Aided, in part, by USPHS-MH 11380.

² Professor of Psychiatry, New York Medical College and Director of Clinical Services, New York Medical-Metropolitan Hospital Medical Center.

patients. The toposcopic variability of rhythms has been shown to be poorly synchronized, and phase reversals and delays in the spread of signals have been described. No identifiable set of characteristics, however, has been consistently related to a psychiatric diagnostic group nor to specific aspects of psychopathology³ (17).

For example, in our analyses of resting EEG patterns in a mixed group of schizophrenic and depressive subjects, we observed that the schizophrenic group had a higher proportion of theta-delta activity than the depressive patients—an observation which contrasts with other reports (5). When the age of the subgroups was considered, the schizophrenic population was found to be 20 years younger on the average than the depressive group, and included a large proportion of young male psychotic patients who had been classified as childhood schizophrenia (16). Since age is a significant factor in the developmental changes in the EEG, we concluded that the differences in EEG characteristics were probably related to age differences in the two diagnostic groups, rather than a behavioral criterion; although the possibility of a subgroup with increased slow wave and spike activity weighting the discrimination could not be excluded.

In studies of EEG responsivity to a pharmacological agent, the selection of the response criterion and the stressor has depended largely on the practical features of safety, ease of application, and reliability on retesting, rather than neurophysiological theories. Many studies have been devoted to the effect of barbiturates on the frontal EEG. The principal measurement has been of fast wave (beta) activity, defined usually as the percent-time or amplitude of frequencies between 18 and 25 c.p.s.

³ This failure may not wholly be ascribed to the techniques of the neurophysiologist for their tools have become increasingly precise. Considering the poverty of the present behavioral classifications, it is equally probable that the lack of an objective behavioral criterion may have obscured any neurophysiological-behavioral relationships that may exist.

The sedation threshold requires the administration of amobarbital on a body-weight basis at a set rate. The observer determines that the threshold has been exceeded by slurring of speech or the appearance of sleep or nystagmus. The beta EEG response is then measured, either with an electronic frequency analyzer or visually, and a curve of the rate of beta activity development prepared. The inflection point on this curve is defined as the sedation threshold (3).

In the pentothal threshold technique, 100 mg. pentothal is administered rapidly, and repeated at 2-minute intervals for three injections (4). The amount of beta spindling and beta bursts, their duration, and the voltage of theta activity define the pentothal threshold. In another technique 4 mg./kg. pentothal is administered at a graduated rate in 2 minutes (5). The onset and duration of beta spindling, delta activation and postnarcotic sleep stages are each measured, and their relationship to standardized curves determined.

While each technique has been related to clinical diagnosis by its authors, none has received systematic confirmation. As prognostic indices, however, the barbiturate response measures have shown consistent changes during various treatments. An early application was the enhancement of slow wave activity in induced convulsions. The development of high voltage slow wave (delta) activity with pentothal during the first week of treatment was related to behavioral improvement and the lack of slow wave enhancement was a poor prognostic sign (6). With psychotropic drugs, treatments which reduced the beta response have been shown to be clinically useful (3, 4); and the reduction in beta activity and the inhibition of a postnarcotic sleep stage has been associated with better therapeutic results than the enhancement of such response (5).

The lack of general clinical acceptance has resulted, in part, from technical problems. The rate of development of beta

activity is influenced not only by the subject's clinical state but the characteristics of the resting electroencephalogram and especially the amount of alpha activity. When tests are repeated on different occasions, the results show considerable variability. The criteria of the measurement are selected because of technical simplicity rather than theoretical grounds. For example, in the sedation threshold such different end-points have been reported as nystagmus, onset of sleep, cessation of verbalization, and the maximum amount of beta activity.

The failure of these tests as classificatory devices may also be related to the choice of a stressor. The selection of barbiturates has no theoretical foundation in central nervous system chemistry, but is based on its characteristic EEG response, ease of measurement, and relation to sleep mechanisms which are frequently altered in psychotic states. Responsivity may better be measured by the use of other stressors, as those more related to the neurohumors active at central synapses—as these are more likely to be related to the basis for the therapeutic response in somatic treatments and to the etiology of psychoses (1, 19). Thus, the neurophysiological responses to stressors related to cholinergic mechanisms as atropine, Ditran, and methacholine; and to adrenergic mechanisms as epinephrine and LSD have provided the best classifications as in the methacholine-epinephrine blood pressure response (2).

In addition to technical difficulties inherent in the variability of biological systems, and the problems of selection of a stressor, there is the problem of selection of a criterion. While the electroencephalogram is the best single continuous measurable index of brain function, the regional variability of brain structures, and the gross nature of available recording and macro-analytic systems make the likelihood of the selection of a suitable criterion small indeed. The quantification techniques in

electroencephalography developed during the past few years using both analog and digital computer methods provide many measures for study. Recent interest has focussed on the problem of averaging the evoked response to a repetitive sensory stimulus (8). Unfortunately, the lack of sensitivity to compounds effecting neurohumors and its high variability make it unlikely that this technique will satisfactorily replace the variety of patterns and greater responsivity of the scalp recorded EEG (25). Yet, as the relation between acute changes in brain wave patterns and changes in behavior is high, continued studies of various aspects of the scalp record EEG are warranted (13). In addition, behavioral responses to a stressor has recently been suggested as an interesting strategy to complement the EEG studies.⁴

Summary

Neurophysiological response measures provide objective measures to classify psychiatric populations according to differences in biochemical organization of the central nervous system. Insofar as the phenotypic patterns of mental illness may be rooted in dysfunction affecting neurohumors, stressors interacting with these chemical substrates hold the most promise. The many quantitative measures of the

⁴ In studies of the electroshock process, differences in EEG response patterns were defined (20) and related to the behavioral response patterns. It was of special interest that the behavioral changes to the stress of repeated convulsions also could be classified into identifiable phenotypic patterns. Thus, four principal patterns of behavioral response were seen, including the patterns of euphoric denial, somatization, panic, and paranoia (21). In assessing the groups who developed these individual characteristics, it was observed that patients with euphoric denial had certain psychological and physiological characteristics in common. They tended to be older, foreign born, had high scores on the California F scale, made more errors on the Gottschaldt Hidden Figures Test, presented greater color and fewer human movement responses on the Rorschach, and exhibited greater amounts of alpha activity in the resting EEG than populations with the other three patterns of response (22). While the individual characteristics for each of the other patterns were not as clear, some criteria were defined. These studies led to the concept that behavioral responses to a physiological stress may also be a useful strategy in identifying homogeneous populations for clinical research and for etiological studies and led to studies of phenothiazines (23) and imipramine (24) as stressors.

EEG response provide tools for the categorization of the response.

Response studies with barbiturate stressors provide approximations to subpopulations, but unresolved technical problems of identification of characteristic responses, the use of unitary indices, and a lack of relationship to neurohumors contribute to lowering the usefulness of these measures in diagnosis. Despite many difficulties, reactivity to barbiturate still has utility as a prognostic measure in somatic therapies.

Further responsivity studies in the development of objective classifications are warranted following the theoretic construct that neurophysiological dysfunction may underlie the clinical differences in the functional psychoses, focusing on the problems of identifying stressors with more specific neurohumoral effects.

REFERENCES

(1) KETY, S. S. "Biochemical Theories of Schizophrenia." *Int. J. Psychiat.*, 1965, 1: 409-446.

(2) For recent reviews summarizing these studies, see ROSE, J. T. "The Funkenstein Test—A Review of the Literature." *Acta Psychiat. Scand.*, 1962, 38: 124-153; and the following additional references not included in this review: FEINBERG, I. "Current Status of the Funkenstein Test." *AMA Arch. Neurol. Psychiat.*, 1958, 80: 488-501; HIRSCHSTEIN, R. "The Significance of Characteristic Autonomic Nervous System Responses in the Adjustment, Change, and Outcome of Schizophrenia." *J. Nerv. Ment. Dis.*, 1955, 122: 254-262; GELLHORN, E. "Analysis of Autonomic Hypothalamic Functions in the Intact Organism." *Neurology*, 1956, 6: 335-343.

(3) SHAGASS, C. "The Sedation Threshold. A Method for Estimating Tension in Psychiatric Patients." *Electroenceph. Clin. Neurophysiol.*, 1954, 6: 221-233; SHAGASS, C. "Sedation Threshold. A Neurophysiological Tool for Psychosomatic Research." *Psychosom. Med.*, 1956, 18: 410-419; SHAGASS, C. "A Measurable Neurophysiological Factor of Psychiatric Significance." *Electroenceph. Clin. Neurophysiol.*, 1957, 9: 101-108; SHAGASS, C., and JONES, A. L. "A Neurophysiological Test for Psychiatric Diagnosis: Results in 750 Patients." *Amer. J. Psychiat.*, 1958, 114: 1002-1010; SHAGASS, C., and KERENYI, A. B. "The 'Sleep' Threshold. A Simple Form of the Sedation Threshold for Clinical Use." *Canad. Psychiat. Ass. J.*, 1958, 3: 101-109; SHAGASS, C., and KERENYI, A. B. "'Sleep' Threshold Techniques." *Psychiat. Res. Rep. Amer. Psychiat.*

Ass., 1959, No. 11: 59-65; SHAGASS, C., MIHALIK, J., and JONES, A. L. "Clinical Psychiatric Studies Using the Sedation Threshold." *J. Psychosom. Res.*, 1957, 2: 45-55; SHAGASS, C., MULLER, K., and ACOSTA, H. B. "The Pentothal 'Sleep' Threshold as an Indicator of Affective Change." *J. Psychosom. Res.*, 1959, 3: 253-270; SHAGASS, C., and NAIMAN, J. "The Sedation Threshold as an Objective Index of Manifest Anxiety in Psychoneurosis." *J. Psychosom. Res.*, 1956, 1: 49-57; SHAGASS, C., NAIMAN, J., and MIHALIK, J. "An Objective Test Which Differentiates Between Neurotic and Psychotic Depression." *Arch. Neurol. Psychiat. (Chic.)*, 1956, 75: 461-471.

(4) GOLDMAN, D. "Specific Electroencephalographic Changes With Pentothal Activation in Psychotic States." *Electroenceph. Clin. Neurophysiol.*, 1959, 11: 657-667; GOLDMAN, D. "Differential Response to Drugs Useful in Treatment of Psychoses Revealed by Pentothal-Activated EEG." In J. WORTIS (editor), *Recent Advances in Biological Psychiatry*, Grune and Stratton, New York, 1960: 250-267; GOLDMAN, D. "Electroencephalographic Changes Brought to Light Under Pentothal Activation in Psychotic (Schizophrenic) Patients, With Particular Reference to Changes Produced by Pharmacologic Agents." *Ann. N. Y. Acad. Sci.*, 1962, 96: 356-374; GOLDMAN, D. "An EEG Correlate of Severe Organic Cerebral Deficiency." *Electroenceph. Clin. Neurophysiol.*, 1963, 15: 157; GOLDMAN, D. "Electroencephalographic Manifestations Associated With Psychotic Illness: Pentothal Activation Technique and Pharmacologic Interrelationships." *Comprehens. Psychiat.*, 1964, 5: 80-92.

(5) ITIL, T. "Elektroencephalographische Studien bei Psychosen und Psychotropen Medikamenten." Ahmet Sait Matbaasi, Istanbul, 1964, 128 pp.; ITIL, T. "Relation of Resting and Pentothal-Activated EEG to Drug Therapy in Depressive Patients." In J. O. COLE and J. R. WITTENBORN (eds.), *Pharmacotherapy of Depression*, C. C. Thomas, Springfield, Ill., 1966: 112-126.

(6) ROTH, M. "Changes in the EEG Under Barbiturate Anaesthesia Produced by Electroconvulsive Treatment and Their Significance for the Theory of ECT Action." *Electroenceph. Clin. Neurophysiol.*, 1951, 3: 261-280; ROTH, M., KAY, D. W. K., SHAW, J., and GREEN, J. "Prognosis and Pentothal-Induced Electroencephalographic Changes in Electroconvulsive Treatment." *Electroenceph. Clin. Neurophysiol.*, 1957, 9: 225-237.

(7) MONROE, R. R. "Episodic Behavioral Disorders—Schizophrenia or Epilepsy." *Arch. Gen. Psychiat.*, 1959, 1: 205-214; MONROE, R. R., HEATH, R. G., MILLER, W., and FONTANA, C. "EEG Activation With Chloralose." *Electroenceph. Clin. Neurophysiol.*, 1956, 8: 279-287; MONROE, R., JACOBSON, G., and ERVIN, F. "Activation of Psychosis by a Combination of Scopolamine and Alpha-Chloralose." *Arch. Neurol. Psychiat. (Chic.)*,

1956, 76: 536-548; MONROE, R. R., KRAMER, M. D., GOULDING, R., and WISE, S. "EEG Activation of Patients Receiving Phenothiazines and Chlordiazepoxide." *J. Nerv. Ment. Dis.*, 1965, 141: 100-107.

(8) JONES, R. T., BLACKER, K. H., CALLAWAY, E., III, and LAYNE, R. S. "The Auditory Evoked Response as a Diagnostic and Prognostic Measure in Schizophrenia." *Amer. J. Psychiat.*, 1965, 122: 13-41; SCHWARTZ, M., and SHAGASS, C. "Evoked Potential Studies in Psychiatric Patients." *Ann. N. Y. Acad. Sci.*, 1964, 112: 526-542; SCHWARTZ, M., and SHAGASS, C. "Recovery Functions of Human Somatosensory and Visual Evoked Potentials." *Ann. N.Y. Acad. Sci.*, 1964, 112: 510-525; SHAGASS, C. "A Neurophysiological Approach to Perceptual Psychopathology." In P. H. HOCH and J. ZUBIN (editors), *Psychopathology of Perception*. Grune & Stratton, Inc., New York, 1965: 41-61; SHAGASS, C., and SCHWARTZ, M. "Evoked Cortical Potentials and Sensation in Man." *J. Neuropsychiat.*, 1961, 2: 262-270; SHAGASS, C., and SCHWARTZ, M. "Cerebral Cortical Reactivity in Psychotic Depressions." *Arch. Gen. Psychiat.*, 1962, 6: 235-242; SHAGASS, C., and SCHWARTZ, M. "Visual Cerebral Evoked Response Characteristics in a Psychiatric Population." *Amer. J. Psychiat.*, 1965, 121: 979-987; SHEPPARD, C., DANIELSON, D., and MERLIS, S. "The Effects of Drugs on the Amplitude of the Visual Evoked Potential: An Exploratory Study." (Personal communication.)

(9) RUBIN, L. S. "Pupillary Reactivity as a Measure of Adrenergic-Cholinergic Mechanisms in the Study of Psychotic Behavior." *J. Nerv. Ment. Dis.*, 1960, 130: 386-399; RUBIN, L. S. "Pupillary Reactivity as a Measure of Autonomic Balance in the Study of Psychotic Behavior: A Rational Approach to Chemotherapy." *Trans. N. Y. Acad. Sci.*, 1960, 22: 509-518.

(10) EYSENCK, H. J. "Conditioning and Personality." *Brit. J. Psychiat.*, 1962, 53: 299-305; EYSENCK, H. J. "Extraversion and the Acquisition of Eyeblink and GSR Conditioned Responses." *Psychol. Bull.*, 1965, 63: 258-270.

(11) POLLACK, M., KARP, E., BELMONT, I., WILLNER, A., and FINK, M. "Comparative Studies of Chlorpromazine and Imipramine II: Psychological Performance Profiles." In P. B. BRADLEY, F. FLÜGEL, and P. H. HOCH (editors), *Neuro-Psychopharmacology*. Elsevier, Amsterdam, 1964, 3: 373-376; LANDIS, C. "Changes in Flicker Fusion Threshold and Choice Reaction Time Induced by ECT." *J. Psychol.*, 1956, 41: 61-80.

(12) KRISTOFFERSON, A. B., and CORMACK, R. H. "Preliminary Results Concerning the Galvanic Skin Response as an Indicator of Drug Effects." In L. UHR and J. G. MILLER (Editors), *Drugs and Behavior*. John Wiley & Sons, New York, 1960: 405-408.

(13) FINK, M. "A Selected Bibliography of Electroencephalography in Human Psychopharmacology 1951-1962." *Electroenceph. Clin. Neurophysiol.*, 1964, Suppl. 23, 68 pp.; WIKLER, A. "Clinical and Electroencephalographic Studies on the Effects of Mescaline, N-Allylnormorphine and Morphine in Man." *J. Nerv. Ment. Dis.*, 1954, 120: 157-175; WIKLER, A. "The Relation of Psychiatry to Pharmacology." *American Society for Pharmacology and Experimental Therapeutics*. Williams & Wilkins Co., Baltimore, 1957.

(14) ZUBEK, J. P. "Behavioral and EEG Changes After 14 Days of Perceptual Deprivation." *Psychon. Sci.*, 1964, 1: 57-58; ZUBEK, J. P., and WELCH, G. "Electroencephalographic Changes After Prolonged Sensory and Perceptual Deprivation." *Science*, 1963, 139: 1209-1210; ZUBEK, J. P., and WILGOSH, L. "Prolonged Immobilization of the Body: Changes in Performance and in the Electroencephalogram." *Science*, 1963, 140: 306-308.

(15) WOLSTENHOLME, G. E. W., and O'CONNOR, M. (editors). "Ciba Foundation Symposium on the Nature of Sleep." Little, Brown & Co., Boston, 1961.

(16) FINK, M., ITIL, T., and CLYDE, D. "Classification of Psychoses by Quantitative EEG Measures." In J. WORTIS (editor), *Recent Advances in Biological Psychiatry*, vol. 8, Plenum Press, New York, 1966: 305-312.

(17) For a full list of citations see reference 16, and: BEKHTEREVA, N. P. "Biopotentials of Cerebral Hemispheres in Brain Tumors." In J. WORTIS (editor), *The International Behavioral Sciences Series*. Consultants Bureau, New York, 1962, 183 pp.; WILSON, W. P. (editor), "Applications of Electroencephalography in Psychiatry." A symposium. Duke University Press, Durham, N.C., 1965: 123-184; SHAGASS, C. "Electrophysiology of Depression" in J. O. COLE and J. R. WITTENBORN (eds.), *Pharmacotherapy of Depression*, C. C. Thomas, Springfield, Ill., 1966: 91-111.

(18) ACKNER, B., and PAMPIGLIONE, G. "An Evaluation of the Sedation Threshold Test." *J. Psychosom. Res.*, 1959, 3: 271-281; BOUDREAU, D. "Evaluation of the Sedation Threshold Test." *Arch. Neurol. Psychiat. (Chic.)*, 1958, 80: 771-775; GOLDMAN, D. Review of ITIL, T. M., "Elektroencephalographische Studien bei endogenen Psychosen und deren Behandlung mit psychotropen Medikamenten unter besonderer Berücksichtigung des Pentothal-Elektroencephalogramms." Ahmet Sait Matbaasi, Istanbul, 1964, 128 pp. In Book Reviews, *Electroenceph. Clin. Neurophysiol.*, 1965, 19: 426-427; ROGINA, V., and KORBAR, M. "Pentothal Induced EEG Changes in Schizophrenic Patients." Paper presented at 6th International Congress of Electroencephalography and Clinical Neurophysiology, Vienna, September 5-10, 1965; SILA, B.,

MOWRER, M., ULETT, G., and JOHNSON, M. "The Differentiation of Psychiatric Patients by EEG Changes After Sodium Pentothal." In J. WORTIS (editor), *Recent Advances in Biological Psychiatry*. Grune & Stratton, Inc., New York, 1962, 4: 191-203.

(19) FINK, M. "Effect of Diethazine on EEG and Significance for Theory of Convulsive Therapy." *Electroenceph. Clin. Neurophysiol.*, 1958, 10: 207-208; FINK, M. "Effect of Anticholinergic Compounds on Postconvulsive Electroencephalogram and Behavior of Psychiatric Patients." *Electroenceph. Clin. Neurophysiol.*, 1960, 12: 359-369; FINK, M. "Cholinergic Aspects of Convulsive Therapy." *J. Nerv. Ment. Dis.*, 1966, 142: 475-484.

(20) FINK, M., and KAHN, R. L. "Relation of EEG Delta Activity to Behavioral Response in Electroshock: Quantitative Serial Studies." *Arch. Neurol. Psychiat. (Chic.)*, 1956, 78: 516-525.

(21) FINK, M., and KAHN, R. L. "Behavioral Patterns in Convulsive Therapy." *Arch. Gen. Psychiat.*, 1961, 5: 30-36.

(22) FINK, M., KAHN, R. L., and POLLACK, M. "Psychological Factors Affecting Individual Differences in Behavioral Response to Convulsive Therapy." *J. Nerv. Ment. Dis.*, 1959, 128: 243-248; KAHN, R. L., and FINK, M. "Personality Factors in Behavioral Responses to Electroshock Therapy." *J. Neuropsychiat.*, 1959, 1: 45-49; KAHN, R. L., and FINK, M. "Prognostic Value of Rorschach Criteria in Clinical Response to Convulsive Therapy." *J. Neuropsychiat.*, 1960, 1: 242-245; KAHN, R. L., POLLACK, M., and FINK, M. "Sociopsychologic Aspects of Psychiatric Treatment in a Voluntary Mental Hospital: Duration of Hospitalization, Discharge Ratings and Diagnosis." *Arch. Gen. Psychiat.*, 1959, 1: 565-574; KAHN, R. L., POLLACK, M., and FINK, M. "Social Attitude (California F Scale) and Convulsive Therapy." *J. Nerv. Ment. Dis.*, 1960, 130: 187-192; KAHN, R. L., POLLACK, M., and FINK, M. "Figure-Ground Discrimination After Induced Altered Brain Function." *Arch. Neurol. (Chic.)*, 1960, 2: 547-551.

(23) KLEIN, D. F., and FINK, M. "Psychiatric Reaction Patterns to Imipramine (Tofranil)." *Amer. J. Psychiat.*, 1962, 119: 432-438.

(24) KLEIN, D. F., and FINK, M. "Behavioral Reaction Patterns With Phenothiazines." *Arch. Gen. Psychiat.*, 1962, 7: 449-459.

(25) BRAZIER, M. A. B. "The Application of Computers to Electroencephalography." In R. W. STACY and B. D. WAXMAN (editors), *Computers in Biomedical Research*, Academic Press, Inc., New York, 1965, 1: 295-315; BURCH, N. R. "Automatic Analysis of the Electroencephalogram: A Review and Classification of Systems." *Electroenceph. Clin. Neurophysiol.*, 1959, 11: 827-834; FINK, M., and SHAPIRO, D. "Analysis of EEG by Digital Computer Methods and Applications to Psychopharmacology."

7th IBM Medical Symposium. IBM Press, Poughkeepsie, N.Y., 1967, 7: 263-290; FINK, M., SHAPIRO, D., HICKMAN, C. and ITIL, T. *idem.* In N. S. KLINE and E. LASKA (editors), *Computers in Psychiatry*, Grune and Stratton, N.Y. (In press.)

Discussion

GEORGE KELLY, Ph. D., *Professor of Psychology, Brandeis University, Waltham, Mass., Chairman*

(Papers: "Typology of Schizophrenia Based on Multidisciplinary Observational Vectors," Nils B. Mattsson, LL.M., Biometric Laboratory, George Washington University, Washington, D.C., and Ralph W. Gerard, M.D., Dean of the Graduate Division, University of California, Irvine. "Neurophysiological Response Strategies in the Classification of Mental Illness," Max Fink, M.D., Director, Missouri Institute of Psychiatry, University of Missouri School of Medicine.)

FORMAL DISCUSSION OF PAPERS

	Page
EUGENE LASKA, Ph. D.	540
MAX HAMILTON, M.D.	542

OPEN DISCUSSION

DRS. MURPHY, PASAMANICK, GREENHOUSE, KRAMER, MATTSSON, KATZ, FORGY, FINK, and GERARD	543
--	-----

Discussant's Remarks

EUGENE LASKA, Ph. D., *Director, Computer Department, Rockland State Hospital, Orangeburg, N.Y.*

The paper by Drs. Mattsson and Gerard is, I think, a heroic effort to use implements from the box of statistical tools even if the "lug wrench" doesn't fit the "nut." Their second sentence says, "*objective*"¹ typologies or classifications of schizo-

¹ *Italic is my own.*

phrenia are of the utmost importance as regards questions about etiology, treatment, and prognosis."

The almost mandatory universal agreement with this statement delineates what most of us would consider an appropriate line of research. Let us briefly review the methodology of this paper to see if *objectivity* is maintained.

Two hundred eight subjects in the main study were observed on 388 variables fairly evenly distributed between the behavioral and biological areas. In order to classify schizophrenia, by way of statistical techniques (well known to be objective), the authors made their first nonobjective decision which was to use component analysis (which the paper states was originally used for such computations by one of the authors). We will call that nonobjective, or more formally, "Arbitrary Decision Zero."

Of the 388 variables, a subset of 108 variables were selected for statistical manipulation. This is "Arbitrary Decision One," the following comment by the authors notwithstanding. "The selection was prompted by the need to rid the matrix of excess redundancy and to bring balance to the representation of behavioral and biological areas."

It should be noted that if one more variable of the 388 were included in the study it would add an additional 109 correlation coefficients to be estimated. Having selected 108 parameters, there are a total of 5,778 correlation coefficients to be estimated based upon the 147 patients having "complete data."

"Arbitrary Decision Two" was to collect only those factors for which the values of the associated eigen-values were greater than 2. The author states that usually 1 is used, but due to the size of the large matrix, the criterion was changed from 1 to 2; this procedure resulted in 11 factors.

"Arbitrary Decision Three" resulted from observing two rotations, one the

Quartimax and the other the Varimax, and adopting one of them because "a clearer pattern" emerged with it.

"Arbitrary Decision Four" was to take a cutoff point for the factor loading of plus or minus 0.30. Thus each patient was "typed" from this arbitrary cutoff point.

The authors go on to describe the characteristics of the resulting 11 groups, a process equivalent to the usual psychological procedure of labeling or naming the factors. This seems strictly a clinical procedure based on examination of the patients' factor scores and test profiles. We won't label it as arbitrary but it should be noted that the authors seem to feel satisfied that the resulting types reflect, "many popular views about schizophrenia." The remarks made by earlier speakers about internal validity appear pertinent here. It is not necessarily true that an analysis is good just because it agrees with what we already previously believe.

Three other analyses were performed but these did not reproduce the 11 types and so they were discarded; presumably for the reason that they didn't agree with the original analysis which agreed with previous belief.

"Arbitrary Decision Five" was to select a subset of 24 from the 108 original variables. These were chosen so as to (a) keep the balance between biological and behavioral variables equal, (b) keep variables which could discriminate schizophrenia, and (c) to represent the stronger test factors in their areas.

Although the third reason seems unclear, the selection of 24 of the 128 variables was the "Fifth Arbitrary Decision."

Subsequent analyses included clustering techniques where a "Sixth Arbitrary Decision" was made as to how to discretize certain kinds of variables. This analysis was deemed unsuccessful because "it yielded five clusters which did not resemble any of the previously obtained patient types" and so the method was dropped.

The "Seventh Arbitrary Decision" occurred in a Q-factor analysis relating to fixing a cutoff point of the correlation coefficient at 0.40.

The "Eighth Arbitrary Decision" is captured in the sentence: "For further statistical workup redundant variables which were useful at the sorting stage were dropped."

There were, at a minimum, eight arbitrary decisions leading to classifications but these did not deter the authors from the conclusion that: "This division * * *, was a result of statistical sorting of objective data, not a consequence of clinical diagnosis."

A set of linear discriminatory functions misclassified only 13 percent of 80 schizophrenics and the authors say that "this result is very satisfactory." It should be pointed out that this 13 percent misclassification was on the very data from which the discriminate functions were estimated.

Finally, the authors and I agree, in their conclusion section, that there are some definite limitations to this study, which appears to be an exercise in making use of statistical techniques (even when they are perhaps not suited for the admirable goals which have been stated).

Discussant's Remarks

MAX HAMILTON, M.D., *Nuffield Professor of Psychiatry, The University of Leeds, Leeds, England.*

Dr. Fink's paper is already a review and therefore there seems very little point in reviewing it further. What he has done is to describe the work in this field, and he has shown, in general, two results. The first is that there is very little link between physiological data, as found in these various investigations, and abnormal mental functioning. The second is that there is not much link with the outcome of treat-

ment. It is singularly unfortunate that three of the best results he reports have not yet even been published. But it is important to recognize that work in this field is more difficult, expensive, and complicated than in almost any other branch of psychiatric research; in this respect, the work of Dr. Fink and his colleagues is almost unique. There are not many workers in this field indeed, and we must not therefore be disappointed if nothing remarkable has yet appeared. It is a new field in which the techniques used are still being developed and we do not yet know in which way new developments may affect the results.

Coming to the paper by Mattsson and Gerard, this consists essentially of a cluster analysis based on variables derived from three fields of study. The importance of this work is that the clusters found from one type of data corresponded with clusters found in another type of data. The groups of clusters thus found are very complex, and I regret to say I find them very difficult to understand. Nevertheless, to me the combination of these three types of variables is astonishingly dramatic and almost unbelievable.

We have already heard criticisms of this work and I am not going to repeat them. The authors themselves have not claimed that it is final or definitive. They have, indeed, asked that this should be replicated. I have very little doubt that sooner or later they will replicate it and I hope that others will do so as well. One point which was not mentioned by the previous speaker, is the nature of some of these variables. Some of them, for example, are derived from the Funkenstein test. Having myself worked in this field I concluded finally that however many stones I lifted up, I would find nothing but horrible wriggly things. One of the difficulties in this type of exploratory research is that one works with material from which one gets results and then one asks one's self, what on earth does it mean? Now the work on

the Funkenstein test has been shown by many people to be meaningless. Here, it would appear for the first time, is a suggestion that it has some meaning. I find this very surprising but perhaps it can be regarded as hopeful.

I would now like to return, as my final point, to the physiological studies described by Dr. Fink. There are some very fundamental difficulties in this work which is not apparent in direct psychiatric observation or psychological testing. Physiological measures and biochemical ones, too, are related in the first place to the constants of the person, in the second place to the constants of the illness and in the third, to functional changes which are occurring at the time. This leads to the problem of disentanglement and it is these three aspects which confuse the issue and which make it difficult to get clear or positive results. These physiological responses are part of the reactivity, the basis of the reactivity, of the individual patient; but here we have to remember that the illness of the patient has psychological and sociological aspects as well, and these cannot manifest themselves in these physiological findings except in the most indirect manner.

The work, therefore, must in the end be linked up with these variables and it is in this direction that one may hope to find something clearer and more specific.

OPEN DISCUSSION

(Dr. Kelly acted as moderator of the open discussion in which conference members participated.)

Dr. MURPHY. The factor of *too little*¹ and *too much*¹ both being related to interference with adaptation applies to a whole lot of variables in the individual, and also to the quantity and quality of stimulation from the environment.

For instance, starting with infancy and early childhood, too low activity level or too high activity level will interfere with

good adaptation; too low sensitivity or too high sensitivity; too low flexibility or too high flexibility, and so on.

Now I feel rather like the child who said, "But the Emperor has no clothes," because as soon as these things are said they are obvious to all of us.

But there are many variables that function this way when we are working on correlations with coping capacity and inadequacies.

Another is not a statistical point: I found myself surprised to find little attention in this meeting paid to the organism and what happens to the organism as a result of experience. This morning there has been more. Here I go back to our own material on infancy. It is possible to recognize basic zones of disintegrative reaction in infancy, and we think of the total of these as contributing to the infant's degree and pattern of vulnerability.

I hope later Dr. Pasamanick will discuss some of his own work and its implications for what has been talked about here. Lauretta Bender's and Greenacre's work are also of interest.

We should be concerned with genetic factors in vulnerability interacting with other factors; the adaptational level of the infant is an outcome of the relation between the early stress and the coping and support contributions from the environment and also from the infant, himself. We can see an emergent adaptational level and also style, with problems in certain zones handled in certain ways.

As the child develops, we find successive efforts toward achievement, mastery, compensation, and so forth, leading to a changed adaptational level. Covert residual vulnerabilities may accompany each successive combination of stress and support. At prepuberty, the children rated earlier by Grace Heider as most vulnerable (on the basis of Escalona's infancy data) had the most acute or severe reactions to the complex organismic and environmental stresses of preadolescence.

¹ Author's emphasis.

I would hope in time it would be possible to look at the different variables talked about here in relation to each other and to see the changes in the organism that can happen as a result of experience and development. The combination, I think, needs to be watched.

(Dr. Pasamanick responded at the subsequent discussion period, but his remarks follow here.)

Dr. PASAMANICK. In view of Dr. Murphy's flattering comments on our studies, I might tell you about some of our predictive capacity in terms of diagnostic entities without appearing to make invidious comparisons although perhaps it might be so interpreted. She referred to our work on brain damage in children following complications of pregnancy and prematurity, which are largely socioeconomically determined. We have been doing some validity studies on the graded types of neurologic impairment in infancy when we used the "Gesell Developmental Examination" by seeing these children during their school ages and examining various aspects of neuropsychologic, social, and school functioning.

Briefly, our ability to predict the presence or absence of disability in the school-age child without weighting the findings by socioeconomic status or seizures (the two most important factors), and even on single psychological functions, is about at the 90 percent level.

We have been able to determine or predict no damage in the school-age child by examination during infancy at almost the 100 percent level. At the same time, we have been able to confirm, at least indirectly, what to us seems a fantastic amount of damage during the prenatal period, amounting to approximately 10 percent of minor damage in a fairly representative sample.

One of the horrifying facts, to us, is that in the lower socioeconomic groups—par-

ticularly Negroes—this level of damage goes up to almost 20 percent of a quite large representative sample.

It is pieces of information like this—vulnerability, or lowering of thresholds to difficulties and stresses, to which Dr. Murphy was referring—which I think we have to take into account.

I should go on to indicate that the correlations between infant developmental quotients and school-age tests of intelligence 6 or more years later when adjusted for socioeconomic status and seizures varies from 0.75 to 0.90 depending upon whether or not the infant was found to be impaired during infancy. In contrast to a good deal of the data reported during the conference this gratifyingly nice predictive ability permits us to contemplate going on to develop diagnostic indices for brain damage, intellectual, and other behavioral functioning in children by examining successive groups of school-aged children who had been seen during infancy and sharpening by weighting the diagnostic criteria in each. It is this kind of historical process which medicine has followed through the centuries which conscious effort and new techniques permit us to think of as hurrying the process along at an increased rate.

Dr. GREENHOUSE. There is an issue here I would like to clarify: in substance it comes from the experience of the mathematical statistician. People tend to forget there is information in all data. This information leads to conclusions which themselves lead to generalizations in varying degrees. The important principle even for mathematical statisticians to remember is that even though research does not abide by all the rules of mathematical statistics, it does not automatically follow that the conclusions are wrong. Whether these conclusions are right or wrong should and will only be determined upon further validation and verification.

I'm not sure one should criticize a validation whose purpose is to verify the char-

acteristics of typologies obtained from previous studies by utilizing as hypotheses the properties of the typologies found to exist in those studies. I would like to point out that when a psychopharmacologist or a clinician is presented with a drug from a drug house, the drug house indicates for what purpose that drug is to be used and the verification research is done in order to determine whether the claims made for the drug are obtainable or not obtainable.

Lastly, I would like to bring up Dr. Laska's criticisms of arbitrary decisions made sequentially along the line. I would be willing to surmise the procedure used by Mattsson and Gerard would be consistent with a basic scheme of independent a priori probabilities which I suggest is assumed by all investigators, whether formalized or not.

Dr. KRAMER. To understand the extent to which the classifications presented here are generalizable, it is important to know the characteristics of the population sampled. The population studied was a heterogeneous one consisting of all males 18-50 years of age, 50 percent of whom had been hospitalized for less than a year, 50 percent for more than a year, with maximum length of hospitalization 10 years. The diagnosis of schizophrenia or non-schizophrenia was based on the demand that three psychiatrists independently called a patient schizophrenic or nonschizophrenic.

Patients who were hospitalized for, say, 5-9 years are the residuals of a cohort of patients who were admitted 5-9 years ago. Such patients who have been exposed to the hospital environment for long periods of time may display characteristics that are a result of the different experiences they have had during their hospital life with respect to diet, treatment, intercurrent illnesses, lack of psychosocial stimulation, etc., rather than of their basic psychiatric disorder.

It would be important to relate the classification of patients hospitalized for 5-9 years to their classification at time of admission to the hospital, so as to determine whether the classifications developed will be independent of length of hospitalization. Such facts would provide important guidelines with respect to generalizability of the results from this type of investigation.

Dr. MATTSSON. Dr. Greenhouse pretty well expressed my feelings about this matter: validation, arbitrary decisions, etc. In this type of research one cannot avoid making a lot of arbitrary decisions, and validation in retrospect is often the only validation available. Dr. Gerard already mentioned that ours was exploratory research intended to go on for a long time.

Dr. Laska's critique hinged upon the word "objective." The mention of objective typologies in our paper refers to the use of so-called objective measurements. In the final typing of the research sample all subjective evaluations were dropped. This, of course, could also be called an arbitrary decision which the researcher, nevertheless, has to make.

Another arbitrary decision was to use the (linear) correlation coefficient as an index of agreement between subjects. This may be a hazardous path to take, unless understood correctly. I believe Dr. Torgerson also pointed this out in his paper. When people correlations are obtained the coefficient may change drastically if some scales are reversed or their signs reflected. This is very important to keep in mind in looking for equality of profile shape between two subjects. We could have used a distance function or just straight cross-products but we were more interested in profile shape and interpreted the possible level separations as intensity differences within the types.

At our project we discussed the matter of scale direction and tried to read each scale in a direction most meaningful in regard to the other scales. Naturally, we

had many bipolar scales; most of the biochemical measures, for example, are bipolar. Many of the social history evaluation scales and the psychiatric ratings are essentially bipolar. There is a high and a low end with something considered normal in between. Both extremes often indicate pathology. Some bipolar scales were involved in our typology analyses and our results might have changed had we reversed these scales. But this was the way they made most sense to us.

Dr. Kramer had a question about our research sample. All patients were male and this is mentioned in our paper.

Earlier in this conference there was some talk about nonlinear relationships. We also looked at this aspect and found that between disciplines there was more indication of nonlinear relationships than within. Within the psychological area, especially, the correlations tended toward linearity. In the presence of nonlinearity, all linear correlations are underestimates of the true relationships.

UNIDENTIFIED DISCUSSANT.² I feel something should be said in relation to what Drs. Laska and Greenhouse have said. I am very impressed in this conference with the depth with which the clinicians have dealt with the problem of classification, and also with their bringing up the question of longitudinal data as well.

Why do clinicians have to go through 8 years of grinding out results to find something they knew? Why aren't there clinicians making stronger statements about the patterns they do see, even if they are longitudinal? Why not let them start accumulating data among clinicians and let this be known to the world so people can begin to discern them? Clinicians should not be afraid to be so-called subjective. Everyone is; even the objective methods have so many arbitrary procedures I would rather

deal with the clinician who at least knows he is being subjective, and may be right.

Dr. KATZ. Some of the criticisms in yesterday's discussion which were leveled at the developers of new typologies had to do with the fact that they had sampled only a very limited area of patient behavior or human functioning. Quite a few of the participants, particularly the clinicians, were disturbed with what they considered to be a very narrow view of the concept of psychopathology. I think that most investigators who are attempting to identify or develop typologies start with such limited areas of human functioning simply because the problems in dealing with the kind of data Drs. Gerard and Mattsson have collected are from a technical standpoint so difficult to resolve. Gerard and Mattsson have attempted to analyze and integrate relationships among diverse areas of behavior and human functioning, in addition to considering the longitudinal question. One of the reasons they were asked by the planning group to summarize their study was because their analytic problems appeared to portend the kind of difficulties which will be faced by investigators in future typological investigations. Diverse kinds of data, e.g., psychological and physiological, will eventually have to be brought together to resolve core questions in this area, and when one reviews the methodologic and statistical problems faced by investigators who deal only with limited areas of functioning, e.g., with behavior, then one can appreciate the complexity of Gerard's and Mattsson's problem. Nevertheless, they have, in a sense, pioneered this direction of research, which it appears will eventually involve a large number of investigators. It would help, therefore, if the criticism of their research were in a constructive vein—considering, for example, how the field can profit from this initial attempt to bring together very disparate types of data, and to seek ways of determining whether a typology exists

² The discussant's name was inaudible to the recorder, but the editors recall the quote is substantially correct.

which cuts across psychological and physiological dimensions.

Dr. FORGY. My comment, like Dr. Katz's, concerns the future. It is diffusely aimed at all the people who have presented empirical papers here. When innovations of theory or technique are demonstrable improvements over their predecessors, then a science progresses. When innovations pile up which scientific consumers can't really evaluate or compare, then there is a tendency for the whole pile to slide quietly back into history.

To prevent this from occurring, I would strongly urge that empirical studies involving new typologies begin using other baselines than the number zero. Even though I am a statistician, I don't think that zero is a good characterization of the present state of our art or arts. Conventional diagnoses, predictions by clinicians who actually see the patients, and ordinary linear combinations of measures probably *also*¹ perform well above zero. These would provide much more informative baselines against which to compare new typologies.

Since both clinicians and computers are already involved in these studies, a relatively small additional effort could add a great deal to the likelihood of cumulative progress in years to come.

Dr. FINK. I would like to direct a question to Dr. Gerard. I was very much stimulated by his paradigm of the tadpole and the frog and the toad and the frog, describing so pithily the genotypic and the phenotypic aspects of classifications. I am concerned that Dr. Gerard has approached schizophrenia from its phenotypic aspects only. The references to the Linnean system seem to emphasize the frog-toad or phenotypic aspects, but in reality, the ultimate test and success of the Linnean system lay in its dependence on breeding as the biological basis for similarity.

I would like to ask Dr. Gerard if, in setting up the original program of identification of the schizophrenic subpopulation, was there a biological basis for the classification, and, if so, to what extent do the data permit a testing of it?

In reviewing the physiological responsiveness measures used for classification purposes in the last decade, their failure may be related to the fact that they were selected because they were simple, safe, and available, and not because they served any true biological basis for discrimination.

Dr. GERARD. We are really very grateful to the discussants, even those who didn't like us, because it helped make very clear some of the issues involved. Perhaps I should paraphrase Mark Anthony. I assure you we have been extremely, acutely aware of all the limitations and defects in the study and showing us sweet Caesar's wounds is just telling us that which we ourselves do know.

Dr. Laska pointed out one is making arbitrary decisions in any experimentation. We made decisions on using a hospital population, on the age cutoff, on eliminating females, on the kind of diet, on the criteria for diagnosis, on the time off drugs, and so on and on and on. The important thing, I think, is that these arbitrary decisions were made a priori in all cases *without looking at outcomes and then deciding*¹ that we will do it this way or that way; they were made in the best light of our judgment of what the situation called for. We did not expect the outcomes we got. I was rather surprised when the particular terms used in subdividing schizophrenics turned out to have a good deal of meaning.

The complaints Dr. Hamilton made are perfectly correct. We don't know what these groups mean. We have rather tried to avoid giving them names or interpretations, but, I submit, just because a lot of

¹ Author's emphasis.

¹ Author's emphasis.

apparently unrelated things have come together does not mean that they don't belong there. I am sure no one a priori would expect that some very severe neurological disturbances, marked indigestion, and low blood cells would turn out to be manifestations of a clear clinical entity, pernicious anemia, but they do.

Dr. Fink has asked about theory and phenotypes. I am glad this orientation came in, even belatedly. I think all we have been doing so far is attempting to deal with phenotypes, hoping that they represent genotypes and will ultimately expose the genetic basis for the phenotypes. We eschewed any theory quite deliberately because we wanted to do iterative research and try not to be guided by our prejudices, which we have certainly inherited. Physiological measures we used were ones available nearly 10 years ago. A good deal of our time went into developing new measures, and some of them we think are considerably better than those in the final

matrix. But since many of the patients had been run through before these were available, we could not include them. They would be very important checkup criteria if, for example, the effect of oxygen on visual threshold or major or minor tremors of muscles and a few other things of that kind, blood chemical differences, did turn up to correlate well with these subtypes. This would be very dramatic confirmation.

At the end of all this, we are back to the original problem. I remind you of the theologian who was chairman of a meeting and introduced a philosopher. In introducing him he said, "Of course, a philosopher is like a blind man on a dark night in an unlit cellar looking for a black cat that isn't there." And the philosopher, getting up to speak, said, "I am afraid I must accept this characterization of a philosopher, but I would only point out that under exactly the same circumstances a theologian would produce the cat."

We think we have produced the cat.

PERSPECTIVES ON THE CONFERENCE

Max Hamilton, M.D., Joseph Zubin, Ph. D., C. Rahdakrishna Rao, Sc. D., F.R.S.

A psychiatrist, a psychologist, and a mathematical statistician were asked to survey the conference proceedings as a whole and to attempt to identify the major issues of concern to the participants. Dr. Hamilton and Dr. Zubin had had time to prepare their remarks. Dr. Rao's comments were extemporaneous. In the final chapter two of the editors reflect on how closely the conference fulfilled its aims and what they see as the major research issues for the future.

Perspectives on the Conference

*Max Hamilton, M.D.*¹

At this conference, a number of psychologists have given their opinions on psychiatric classification from the point of view of psychological and sociological theory. Although some of the most scholarly papers at this conference (especially those from the epidemiologists) have given an account of classification in medicine, the full implication of this approach has not been made sufficiently clear. It would therefore be useful to consider some general problems along these lines.

Whereas medicine has always started from the practical aspect (for it is essentially an applied science), psychology and sociology have been largely theoretical in their origin. Their practical applications have come at the end rather than at the beginning of their work. Psychiatry must be regarded as one of the main branches of medicine. It is true that it has links with other sciences such as psychology and sociology, but it is not unique in this respect, for this is equally true of preventive medicine. For this reason, psychiatry has much to learn from the methods of classification used in medicine, and the history of the development of these methods casts a ray of light on the present difficulties.

1. Classification Cannot Be Complete

Completeness implies that the limit of knowledge has been reached, and this is obviously untrue even in general medicine, let alone psychiatry. The last quarter cen-

tury has shown tremendous advances in the development of classification in general medicine. The epilepsies and bleeding diseases have been subdivided with the increase in knowledge; new virus diseases have been demarcated from the general mass of unknown infections.

These examples have been selected with intent. The subdivision of the epilepsies and bleeding diseases has not destroyed the classificatory term epilepsies or bleeding diseases. It has amplified it by subdivision. On the other hand, the demarcation of new virus diseases has meant that of a vast mass of disorders, a group has been cut off and distinguished from the rest.

The theory of autoinimanity has shown linkages between diseases previously regarded as completely different. Many urgent problems still remain, for example, the various demyelinating diseases have been described as syndromes, but the relationship between them is as little understood as is their nature. Much the same can be said for many problems in dermatology.

These examples exactly parallel the problems that have been considered here at this conference. The problem of the unknown was dealt with in Dr. Rao's paper when he emphasized the importance of an *et cetera* category. There must always be a group for the unknowns, the unclassifiable. When a classification finds that 10 percent, 20 percent, or whatever it is of patients, are not classifiable, this is not an adverse criticism of such classification. On the contrary, it is a point in favor of it. If a classi-

¹ Nuffield Professor of Psychiatry, the University of Leeds, Leeds, England.

fication included everybody it would quite certainly be wrong.

2. Changing Trends

It is important to recognize that the methods used for delineating a syndrome may change with the advance of knowledge. Thus pernicious anemia was originally defined in terms of its symptoms. As the methods of examination of blood improved it was then redefined in terms of a high color index and this was later changed to the more critical findings of anisocytosis. Later it was recognized that achlorhydria was a fundamental characteristic of this disorder. In the course of time its relation to subacute combined degeneration of the cord was observed and this connection was explained only by the discovery of the common cause to these two very different disorders. In the same way, coronary thrombosis was first delineated by demonstrating the association between certain symptoms and postmortem findings. Later, the development of the ECG enabled physicians to make the diagnosis during life and even to detect very minor conditions. Nowadays the SGOT test is regarded as of even greater importance than the ECG. In psychiatry, Kraepelin defined dementia praecox in terms of a common outcome arising from certain patterns of symptoms. In time he came to recognize that the so-called dementia was not inevitable. These patterns of symptoms were then further extended by Bleuler to embrace the disorder we now know as schizophrenia, or the group of schizophrenias.

3. Severity

The papers read at this conference (except Dr. Torgerson's) have unfortunately not given sufficient attention to the problems of the differences between a severe and a mild form of one disorder. Not only may a change in severity require a change in disposal of the patient, but it may also determine the use of a different type of

treatment. Dr. Gardner pointed out what may happen if these are confused. Furthermore, as a particular disorder changes from mild to severe, the total number of symptoms may increase and even the kind of symptoms may change. The reverse may be seen when a particular patient recovers and his illness changes from severe to mild. In general medicine this factor has prevented classification from being too closely linked to treatment or the response to treatment or the prognosis. Acute appendicitis remains acute appendicitis whether it is treated by antibiotics or by surgical operation; pulmonary tuberculosis is the same whether the disease is resistant or sensitive to streptomycin; a patient may die from severe pneumonia and may recover from a mild one. The same is true in psychiatry; not all schizophrenics inevitably deteriorate; some patients suffering from depressive states do not respond to electroshock treatment and so on.

4. Pathognomonic Features

A disorder or syndrome is characterized by a particular pattern of symptoms, signs, and other features. The diagnosis becomes more certain as the number and clarity of these features increases, but they are not of equal importance. Some of them are minor or even trivial, others are fundamental. Furthermore, advance of knowledge may alter the relative value of different features of a syndrome. We are now willing to diagnose pernicious anemia in the absence of anemia, but not in the absence of achlorhydria. At one time the latter sign was not even included in the definition of the disorder. The same is true in psychiatry. Not all schizophrenics show hallucinations and delusions; although elation is a central symptom in the syndrome of mania, many manic patients do not show elation. The various symptoms of anxiety are irrelevant for a diagnosis of schizophrenia but they are all-important for a diagnosis of anxiety state. It could be argued that a diagnosis

of depressive state could be justified even in the absence of a depressive mood. The paper by Dr. Klein brings out this point in its emphasis on the setting of symptoms. Dr. Lehmann in his paper has also argued how important this point is. In mathematical terms it may be stated as one aspect of nonlinearity.

There is another aspect to the question of diagnosis. The process of making a diagnosis is a decision tree. The physician starts off with information which leads to certain probabilities for various diagnoses. As the information increases it diminishes some probabilities and increases others but at any one point it is possible for the diagnostician to go back up the tree and start afresh. Many of the techniques described at this conference do not have the decision tree as their fundamental basis.

Before going on to deal with the fundamental questions raised by this symposium I would like to mention one minor matter which has been conspicuous throughout this conference by its absence. I refer to the general disregard of clinical skill. The most delicate instrument for the detection of patterns that we have is the human mind, or if you prefer it, human perception. Even those of us who have only a nodding acquaintance with the visual arts can differentiate quite easily between a painting by Van Gogh and one by Rembrandt, but the expert requires only one glance to enable him to ascribe a date to a painting by Rembrandt. In this conference the technical terms, apathy, withdrawal, and retardation have been used as if they were interchangeable. They are not. Vague terms such as aggression or excitement are used to describe a number of very different patterns of behavior. A layman is entitled to use the word energy in a vague or figurative sense, but this would be unforgivable in an engineer or physicist. At this conference it is reasonable to assume that the imprecision arises from carelessness in writing and speech, but it is a real difficulty in day-to-day clinical work. Any

method of classification in psychiatry will be severely handicapped if psychiatrists do not use technical terms precisely and do not take sufficient care to distinguish between the phenomena which these terms describe. This is even more important in research, for whereas it is always possible to combine data when it has been shown that a distinction has been made when there is no difference, the reverse is not possible. It could be fairly argued that some of the papers presented at this conference do show some faults in this respect.

On this question of clinical skill I would like to raise a little point of which we are all aware but which has never been mentioned. At this conference there are psychiatrists, clinical psychologists, and statisticians, and there is no doubt about the concealed hostility between these three groups, an attitude which is not only mistaken but quite unnecessary. In clinical practice, the physician who treats a patient and has responsibility for him calls upon the biochemist, the hematologist, and the radiologist for technical assistance. He does not despise them, nor does he regard them as mere assistants. They do not feel their status is threatened because he has the responsibility for the patient. They are technologists, sometimes scientists, workers in their own right, and they are collaborating in a particular task, with the contributions they can make.

The same is true in the field in which we are working. The psychiatrist is a technician, a professional man who has a job to do, and his work requires a specific perceptive skill. It is his business to observe, to take note and, if necessary, to interpret; but it is a very limited skill. The ability to interview a patient does not confer an ability to interview a candidate for a commission in the forces. This was very well demonstrated during the war when it was found that the psychiatrists had a negative contribution to officer selection.

There are many kinds of interviews. A psychiatric interview is not the same as a

lawyer's interview, a police or job selection interview. A psychiatrist's skill is not to be despised because it is limited, but it is not to be overrated and thought, by those who do not have this skill, to have some mystical or mysterious quality.

There is no mystery or secret about the clinician at all. He has been trained to do a specific job, so has the psychologist and the statistician. In the work with which we are concerned here, the data is gathered by the observer. If the observer is no good, nobody need bother any further, but what is done with the data and how it is collected are the subject matter for the statisticians and the clinical psychologists, and no progress will be made without them either.

It now remains to consider the questions raised for the final discussion of this conference.

a. How is a Good Typology Validated?

The value of a classification depends on showing that when a certain type of data is subdivided into groups this is equivalent to forming the same groups from additional data. When we classify animals into those with hair and without, we are, at the same time, more or less classifying them into those that have their young born alive as against those not, into those who suckle their young against those who do not, who have a steady temperature and so on. The whole point of the subdivision into mammals and other animals is precisely this: that each type of classification produces much the same grouping.

In psychiatry the main source of data is from the symptoms of the patients, and this term also covers not only the behavior of the patient but also his statements (not necessarily complaints) e.g., what he says about his hallucinations and delusions. We have, in addition, data from family histories, genetic data in the more general sense, information on aetiological factors, and on the course and outcome of the illness. To this must be added the response to treatment and psychological, physiologi-

cal, and biochemical findings. It is obvious that as additional information is obtained there will be minor modifications in the groupings already established on the basis of previous data, but insofar as groupings remain relatively constant the additional data confirm the value of the classification. In this sense validity is identical with value, though of course it has other meanings.

Work along these lines has already been reported in this conference and will continue. It is important to emphasize the sequential procedure outlined here. Classification based on every kind of information lumped together is not as meaningful as the slow accumulation described. The work reported at this conference will have to be proved capable of replication, for in the end, the most important way of convincing users that a classification is valid is to show that it is of clear practical value in organizing the material; this is possible only if it can be repeated without difficulty and continually.

Furthermore, it must have a practical outcome. Psychiatry is a technology; its purpose is to treat patients and to deal with them; in the end, the value of a classification will be determined by the practical test.

Dr. Lehmann has said that the work on drug-responsive groups is of very little value in the clinic and this is largely true. The differences found are indeed very small and are statistically significant only with a sufficiently large number of cases. Nevertheless, we must not confuse the requirements of the clinic with those of research. That which is of little use in clinical practice may give an important lead to further research. This is what I mean by the practical test: the proof of the pudding is in the eating.

b. Clinical requirements of a classification

It is obvious that a classification must be neither too broad nor too narrow. It is of no help to define all our patients as being merely ill since this yields no information.

On the other hand if the classification is too narrow, it not only loses reliability but it can become lost in the trivia of minute details.

One exception must be made to this: Even a rare disorder which is found only once or twice in many thousands, should be distinguished if it can be. In other words, a fine subdivision is not strictly related to a frequency distribution.

The problems which result from a too narrow classification have been well exemplified in the paper by Dr. Weinstock. Speaking personally, his difficulties must be considered to arise from a fundamental incompatibility between modern classification and psychodynamic theory. The latter is a theory of aetiology based on a Hippocratic viewpoint, which originally regarded all disease as due to imbalance of the four humors. Psychoanalytical theory regards all mental disorder as resulting from varying interrelationships between a few psychodynamic processes. Its whole outlook is therefore completely different from that of modern medicine which originated from Sydenham's work in cutting up clinical phenomena into blocks; the syndromes.

Classification must be based on relevant material and unfortunately it must be admitted that at the moment current psychological tests have not yet adequately demonstrated their relevance. As for biochemical and physiological tests, these must still be regarded as completely peripheral and of interest to research workers only.

A classification must be practicable. This means that it will be of little use until it is constructed in such a form that it can have widespread use. From the epidemiological point of view this is of fundamental importance. Nationwide statistics cannot depend upon the findings of a few specialized laboratories. In the first place it must be based on enduring features of the illness rather than on temporary and fluctuating ones, and in the second place the groupings must be identifiable under

ordinary clinical conditions. It should have a value in relation to prognosis and treatment but this is not its sole or even primary aim. Finally, good classification should be clearly seen as an extension or outgrowth of current broad groupings.

A classification should have a clear purpose or purposes, including those at different levels. For example, the clinician is accustomed to make a diagnosis of simple retarded depression, acute depression, and stuporose depression (or depressive stupor). He uses these terms in preference to the more simple mild, moderate, and severe because they are more descriptive. He includes them in the diagnosis because the severity of the illness is of great importance to him. To the statistician or administrator they are all just depressions, and to attempt to introduce them into his classification would probably be unhelpful, if not confusing, since the various grades are ill-defined. This statement must be qualified, for the severity of illness may have significance for the disposal of the patient and other administrative measures. Classification is man-made to serve manifold and hierarchical uses. Classification for different purposes need not be identical but great overlap is desirable.

These requirements show clearly that the work reported at this conference is by no means complete. It will have to be repeated and further extended but, in addition, it must be made clear that the various groupings described should be translated into clinical terms; in the end they must be reduced to clinical descriptions in literary form.

c. What Needs to be Done Now?

It is not for any one person to tell research workers what they should do and, with this proviso, I would like to make the following suggestions. I have already mentioned certain problems and obviously some of them could be dealt with easily. In the first place, it would be necessary to investigate the stability of syndromes de-

rived from factor and cluster analysis. It will be an important step forward if it can be shown that successive assessments, extending over a long period up to many years if necessary, will allocate a given patient into the same group. Another extension of present work, some of which has already been reported here, is to demonstrate the relation between here-and-now syndromes to the course and outcome of treatment.

Furthermore, it will be necessary to compare them, particularly when we are considering the response to treatment, with traditional syndromes. Are they any better? Do they supplement?

A much more difficult problem is that of trying to evaluate a hierarchy of symptoms and other findings. The technique of discriminant functions does give different weights to variables when these are used for discriminating between different pairs of syndromes, but the mathematical model is still based on the linear assumption, as Dr. Cole has pointed out, and this is not yet adequate. Much work will have to be done, both practical and in the development of theoretical models, before psychometric techniques can be regarded as matching the subtlety of human perception and judgment. What is needed is a theory which is sufficiently complex to deal with the material and then the devising of suitable techniques for proving or validating it.

One final point: It is of course impossible to predict the lines of future research, but I have little doubt that the critical and fundamental developments in psychiatry will be based on biochemical and genetic work in the future. But biochemists and geneticists are blind and helpless in an unknown field without the clear guidance of precise clinical observation and description. The phenylketonuric syndrome provides an excellent example. Psychiatrists and psychologists will therefore have an essential role in future developments and this is the task for which they must prepare themselves.

Perspectives on the Conference

*Joseph Zubin, Ph. D.*¹

I. Introduction

The technique of the 10-minute précis and the 2-minute discussion interval has produced a new time unit which makes it essential that thoughts move rapidly and be expressed fast. I shall, therefore, limit my remarks to only four points.

- (a) What were the points of full agreement of this conference?
- (b) What were those in which only partial agreement was attained?
- (c) What were the disagreements?
- (d) What points which should have been included were omitted from our discussions?

Initially I had hoped to make a content analysis of each paper, punch IBM cards for each contributor, and make a statistical analysis of the agreements through either a typological, dimensional, or like-mindedness approach; but unfortunately time ran out.

II. What are the points of agreement?

1. Diagnosis is in a bad way:

Present-day diagnostic categories are too heterogeneous, and, unless improvement is introduced, we will fail to meet the needs of the new types of patients in community mental health centers that are starting to flow into our facilities and demand diagnosis and classification.

2. Typology seems to be more congenial to the usual clinical way of classification, but the burden of proof lies on the typologist to demonstrate that his method has more merit and more power than the clinical or dimensional approach.

¹ Chief of Psychiatric Research (Biometrics), Department of Mental Hygiene, New York, and Professor of Psychology, Columbia University.

III. Points of partial agreement:

1. Diagnosis has a multipurpose goal.
2. Better data than those now provided are needed.
3. Typology tends to reduce intragroup variability. This is a gain for experimental purposes.
4. Prior clinical knowledge could be helpful in the preparation of diagnostic schema. Of course, the difficulty here is that much of prior clinical knowledge "ain't so" and how to distinguish between the reliable knowledge and unreliable knowledge is an important question.
5. Arbitrary decision must be resorted to often in the making of any classification and the nature of this arbitrary decision must be at least specified so that replication is possible.

IV. What are the points of disagreement?

1. Though clinical interviewing is loose and freewheeling, it is not yet clear whether it is inferior to the more systematic structured approaches, like the Mental Status Schedule or the Structured Clinical Interview. Work going on in our own laboratory seems to demonstrate that these techniques are at least more reliable. Whether they are more valid remains to be demonstrated.
2. The relative merits of factor analysis and typological analysis are not yet agreed upon. Back in the late 1930's I wrote a paper entitled "Sociobiological Types and Methods for Their Isolation" and it appears to me that I did better then than I knew by introducing the term sociobiological types. The types that we got then and those we get today are ad hoc types and not biological types. That is why it is important that we look into the social cultural background of the people for whom profiles are developed. High degree of hostility in a person coming from a low socioeconomic level where aggressive behavior and overt violence are more often seen than in other sociocultural groups, must be separated

from a similar score for a person from another socioeconomic level. This may be the reason why our types do not hold up upon reexamination and unless we include the social cultural factors as well as other factors that are necessary, such as educational background, we will not be able to make much progress.

3. The relative merits of disease entities versus behavior patterns is not agreed upon. No definition of disease was arrived at although a rather tentative definition was suggested which runs something like this: A disease is a progressive condition which unless attended to would lead to premature death or to excessive lowering of efficiency. A defect is a stationary condition which leads to an extreme reduction of efficiency. It was pointed out at the conference that such factors as military service or poverty or delinquency or old age would qualify as diseases according to this definition. I doubt this since these terms are too global. Only by analyzing these global concepts into their components can any understanding of their relationship to disease be attained. In the case of old age it is necessary to see what biochemical changes are in back of the gradual lowering of vitality. With regard to poverty, unless we look into such factors as malnutrition or overcrowding which comes with poverty, we do not really deal with a quantifiable or definable concept. Furthermore it is quite clear that only in mathematics can definitions be foolproof and rigid. In biology rigidity of definition falls by the wayside, and the power of the defined concept to integrate observations becomes the criterion of a good definition. Take, for example, the concept of species, which Julian Huxley points out could not be defined rigorously at all. It is a combination of several indicators plus a little flair which makes the definition acceptable. Thus the presence of many open-ended definitions in our field should not be regarded as a hazard.

5. Clinical psychological tests have not fared too well at the conference nor in actual usefulness in everyday work.

6. The medical versus the behavioral model has not yet been discussed fully enough to arrive at any conclusion.

7. In dealing with profiles the distinction between level and pattern has not been duly considered, and it's quite clear that following Penrose one can break up $\overline{d^2}$, the average square deviation, into two components: (1) $(\overline{d})^2$, the square of the average distance between profiles, and (2) s^2 , the variance of these discrepancies. The square of average, $(\overline{d})^2$, of course refers to the discrepancy in level between two profiles and the variance refers to the discrepancy in shape. If complete parallelism exists (all the d 's being constant) the variance drops out completely. To the extent that there is some variance present there is some discrepancy in shape.

8. Specification of the mathematical model underlying clustering met with a good deal of discussion. It is quite clear that some people do not care to specify ahead of time the assumption underlying clustering, and are satisfied with blind clustering techniques. As a result only ad hoc typologies emerge. On the other hand, if you specify normal distributions or rectangular distributions, and specify also the range of correlations between the variables in each cluster, you may be specifying too much. Only by trial and error or by some better systematic approach can this problem be resolved.

9. It was not quite clear as to how one crosses boundaries from the concepts used by one discipline, such as nursing, to another, such as the psychiatric or psychological disciplines. The fact that the same words are used does not necessarily assure us that the concept behind the words is the same.

10. One of the criteria often used by the newer approach to diagnosis is that the obtained results in a cluster or a profile agree

with previous clinical knowledge. Since there is no guarantee that the older clinical formulation is correct, this constitutes a doubtful criterion. In our own work we are most suspicious when a finding corroborates a previously held clinical opinion. Agreement per se is pleasant because it leads to a friendly feeling, but it is no proof.

11. Linear versus nonlinear relationships. It was pointed out that whenever relationships are obtained between variables from two different disciplines—e.g., psychology and physiology, the relationship usually turns out to be nonlinear. Whether this is a result of the units used or whether it represents the influence of more than one underlying factor is an open question. To be sure, not all the correlations within a given discipline are invariably linear—that is why typology is a useful concept.

V. What matters have been either omitted or not discussed sufficiently?

1. One of the major difficulties in the area of improving diagnosis is the lack of criteria for the validation of diagnosis. At the present time there are two aspects that are most urgent, namely the severity of the illness and prognosis of the illness, and these two aspects could be utilized as a basis for validity. At the present time validity criteria for the severity of the illness is available only in the judgment of other clinicians, and utilization of concurrent validity for this purpose is the only way out. With regard to prognosis, the only solution is to have good followup studies. In the meantime the best that one can do with regard to validity is to utilize consensual reliability.

2. The investigation of the decision process involved in diagnosis is of the utmost importance since at the present time the basic tool for the development of diagnosis is the clinician himself. The investigation of this instrument—the clinician—is most appropriate, and methods are

already available for studying decision processes.

3. Insufficient time was allotted to the discussion of the role of personality in its relation to psychopathology. Here the three possible connections between personality and psychopathology ought to be surveyed: (1) psychopathology and personality as one and the same, (2) psychopathology as an interference with personality development, and (3) psychopathology as independent of personality.

Apparently personality assessment, regardless of whether it is independent of psychopathology or dependent on it, is an important element in diagnosis and prognosis.

4. Very little consideration was given, surprisingly, to the distinction between primary and secondary symptoms in such illness as schizophrenia and other illnesses. At one time this was one of the major issues before diagnosticians, but now it seems to have been relegated to the background. It would be interesting to know why that is so.

5. Freud's primary process and secondary process have never been mentioned, and that, too, raises an interesting question.

6. Although much attention was paid to drug therapy in consonance with the nature of the conference, little if any discussion of psychotherapy and the newly developing behavior therapy was engaged in.

7. The mixture problem, of finding the natural lines of cleavage in a heterogeneous population, was left unattended to, probably because of the absence of Professor Barnard.

8. It is interesting to raise the question why the role of psychogenic versus other factors was little heard of in the discussion.

Conclusions:

Conferences such as these, to which we each bring our prejudices and sharpen them on the whetstone of opposition, are not the best way to take diagnosis out of

its present-day doldrums. Because we need better diagnoses today as a result of the tremendous number of therapeutic approaches from which we have to choose, it becomes necessary to take a drastic step. I would like to propose the establishment of assessment centers in various regions of the country where intuitive clinician and objective biometrician confront each other not across the conference table, but in front of living case material or, at least, in front of video tapes of living case material. Only by a direct confrontation with living clinical material can we hope to make progress. Perhaps, following Dr. Klein's suggestion, a combination of assessment centers with community treatment centers is the next step in the development of improved diagnostic schema.

Perspectives on the Conference

C. Rahdakrishna Rao, Sc.D., F.R.S.¹

There has been some discussion during this conference about the role of a statistician in experimental investigations, his role as a collaborator or advisor in scientific research in general. Some have overemphasized his role, and others were somewhat skeptical about it.

I am glad that Dr. Hamilton has just spoken about the usefulness, or the need for collaboration between statisticians and other scientific workers. May I make a few remarks to supplement what Dr. Hamilton has said?

Statisticians often complain that they have to deal with data collected without their advice. The reasons may be many. The data given them are such that they need a complicated analysis for extracting the desired information. Or the data col-

¹ Director, Research and Training School, Indian Statistical Institute, Calcutta, India.

lected do not provide the desired information because of lack of controls and other precautions necessary to eliminate extraneous factors. Or the data are not numerous enough to provide the information with the desired degree of accuracy.

Now I am sure if a very well-defined problem is referred to a statistician and his advice is sought, he will be able to suggest an optimum utilization of the experimental material; that is, he will be able to say what is the most economic way of conducting the experiment to achieve a desired degree of efficiency.

But in such a case he did not have any hand in formulating the problem. He was confronted with a specific problem, and he could give the appropriate advice. He might not have known the reason for any particular experimental material or the treatments chosen by experimenter.

But I believe a statistician can make a more useful contribution in the actual formulation of the problem itself and planning the whole sequence of experiments which the investigation needs. Now this he can do only if he has the chance to understand the subject matter in which the investigation is being conducted. Otherwise his advice would only be superficial in character.

You have to educate a statistician if you want to get something out of him about your own subject, and the object of your investigation—what you are seeking for, the gaps in the existing knowledge, and the nature of the experimental material.

I was somewhat unhappy this morning to listen to Dr. Clyde when he mentioned the problem of determining correlation between two characters. He was specifically referring to the choice of the range of variables which would lead to establishing correlations between two characters.

I think he is in a very dangerous situation. In the diagram he drew, suppose we have height on one axis, and weight on the other. Now we know, that in any homogeneous population, say among the Ameri-

cans, the correlation between height and weight is of the order of 0.2. And this correlation is pretty nearly the same if you study any group of people in India.

But if the object is to show that the correlation between height and weight of individuals is of the order of 0.6 we need only mix the data on samples of Americans and Indians and compute the correlation coefficient. Similarly we can reduce the correlation by mixing the samples from suitably selected populations.

So you have only to ask the scientist who is seeking advice as to what correlations he wants to show in his material [laughter] and you can advise him accordingly.

It is a standard example that I use for my students, to show how spurious correlation can arise due to heterogeneity of data.

So if somebody asks whether a particular technique would yield meaningful results on a given sample there is no easy answer. At least in the problem of height and weight I mentioned we know that the correlation coefficient is meaningful only if the population from which the samples are drawn is homogeneous in a well-defined sense. But when one is dealing with drugs on one axis—I don't understand how drugs of different makes are quantitatively represented—and psychoanalytic responses on the other axis, one is completely in the dark [laughter].

It is an extremely dangerous situation. It may be that these two variables correlate depending upon the nature of the chemicals in the drugs chosen and the nature of the symptoms or measurements that one makes.

As another general remark, I would say there is a tendency on the part of applied workers to use statistical methods somewhat indiscriminately. I am especially referring to work involving cluster analysis, principal component analysis, and so on. They have all been used without yielding much useful results.

Recently, I had the opportunity of reviewing about a hundred papers dealing with the application of principal component analysis, and in most cases the applications were made without any definite purpose. I ended up writing a long paper on this subject discussing the different situations in which the principal component analysis can be profitably undertaken. (Rao (1964) "The Use and Interpretation of Principal Components Analysis in Applied Research," *Sankhya*, 26, 329-358.)

There is also a tendency to use particular models and draw inferences from data by using statistical tools applicable to a chosen model. But such inferences may depend on the validity of the model which may or may not be verifiable from available data. If possible, a hypothetical model intended to be used for data analysis has to be verified on the basis of the available data. Further, it would always be a good invest-

ment to collect data in such a manner, involving extra expenditure if necessary, that they enable us to examine the validity of a set of possible models intended to be used in the analysis of data. As a general rule, in applied work it is necessary to realize that a statistical tool is not to be considered as appropriate for a particular model but more appropriately as an aid in interpreting observed data through different models. (For example, see the discussion on t-test, p. 420 of "Linear Statistical Inference and its Applications" by Rao (1965).)

I hope, I am not drawing a pessimistic picture. I am only suggesting that a good deal of caution is necessary in the application of statistical techniques. Of course, if they are used judiciously and in the right kind of problems, one should expect good results.

Reflections on the Major Conference Issues

Martin M. Katz, Ph.D., and Jonathan O. Cole, M.D.

The planning of a conference can be viewed somewhat like the planning of an experiment. The planners have certain aims, certain major questions to be answered and the conference is structured so as to best achieve these goals. The conference may result in definitive conclusions directly related to the aims—or it may not; it may result, as do most experiments, simply in increased clarification of the major issues—and then, more precisely defined questions—which hopefully set the tone for more definitive work in the future. The major difference between the conference and an experiment is that in the former a much larger number of experimenters are brought together, all with their own personal views of the major obstacles to progress in the field and their own solutions or partial solutions. In this sense a conference, depending on its size and scope, becomes more like an attempt at empirically surveying a field which is currently too broad and too diffuse, and where more pointed interchange and greater understanding among differently oriented people may lead to more rapid progress in the field in general.

The people brought together at this conference represented a very broad range of disciplines and interests—and their differences are reflected in some of the assumptions they had about the nature of the problems with which they were confronted.

In this final chapter the editors will attempt to focus on several of the core

issues which recurred throughout the proceedings. These issues are tied to the major questions raised by the planners and responded to by the participants.

What is the purpose of a classification system? When psychiatrists, clinical psychologists, basic scientists, epidemiologists, and statisticians considered this question, their answers, as might be expected, turned out to be based on quite different assumptions. The clinical psychiatrist, as characterized by Lehmann, is primarily an empiricist and pragmatist; he is concerned with the utility of a classification system either as a practical instrument for assisting him in carrying out his clinical work, or as an applied research procedure for predicting the results of new or old treatments. A new classification system has to be demonstrated to be more helpful in these respects than the current system for him to consider revising his procedures.

The basic scientist in psychopathology is not concerned with the practicality of the classification system but only with its validity. Does a particular system reflect the true state of affairs? Are the mental disorders really made up of these particular configurations of symptoms and characteristics? If the system is valid, and if the characteristics of the treatments which are to be applied are well understood, then it is very likely that the results of a given treatment can be predicted. But it should be clear that the major aim of his research is not to predict the effects of a particular

treatment—but to develop a system which adequately reflects the nature of the phenomena which are being studied.

It was pointed out in the course of the conference that if it is prediction of response to a particular treatment that is really at issue, then the typological approach may be more complicated and less efficient than other approaches. There are ways of quantitatively weighting and combining the available data on patients which may provide more direct solutions than the typological to prediction problems. These more direct procedures do not require an understanding of why a particular combination of variables is effective in predicting the results of treatment (we refer here to the discriminant function and multiple-regression approaches).

Thus, the predictive formulas which are derived from these procedures, require repeated tests of their validity. Because the formulas are so specific in their applications; i.e., they are tied to a particular treatment, new formulas have to be developed and validated for each new treatment, even though the new treatment may be only slightly different from the last one. The approach is pragmatic, it serves a particular purpose and may have little or no generality to other treatments. Yet, as noted, it is the most direct attack on the problem of prediction and with the new technical developments in this area (referred to by Lubin and Klett) it permits the analysis not only of the additive relations among variables, but, also, of certain interactions among the predictor variables. In utilizing the contribution of such interactions to the prediction, these techniques achieve the greatest possible precision and efficiency of prediction when dealing with large samples of patients and large numbers of variables. It is a genuine source of amazement to some in this field that typologies based, as they are, on rather approximate procedures of analysis, procedures which, from a technical standpoint, are so much more vulnerable to criticism than the

multiple-regression approaches, do as well as they do in the area of prediction. The reasons for this are not entirely clear but they have something to do with the fact that to the extent that a typology of the mental disorders validly describes the nature or the structure of the phenomena under study, then predictions concerning the effects of a defined treatment can be developed on a theoretical rather than on a strictly mathematical basis.

The problem of the psychopathologist, however, is not restricted to predicting response to a treatment. He has a more general mission and that is the understanding of the phenomena which make up the substance of mental disorders. His problem is to identify through observation and through theory the more significant or important variables in this field, to develop ways of measuring these variables, and then to investigate the ways in which these variables interrelate. It is possible that these variables do not relate in ways that result in discrete classes or types of people, as was suggested by several of the participants. Torgerson, for example, argued that the underlying structure of mental disorder is still at issue and that the task is to examine this underlying structure before committing the field to the typological or to the dimensional model. It may be that both models are valid; i.e., some of the major variables are continuous and normal in their distributions, others are discontinuous and an analysis of their interrelations results in the emergence of qualitatively distinct classes of subjects.

It was interesting in this respect that the problem of determining whether qualitatively discrete groups, subpopulations or classes, exist within a larger population (otherwise known in statistics as the mixture problem), is still unsolved. The answer to the question raised by Torgerson as to the nature of the underlying structure of mental disorder must, therefore, await the solution of the mixture problem. Needless to say it must await, also, the resolution of

other focal problems in psychopathology, which have more to do with identifying the crucial variables which should enter the analysis and how well they can be measured. Until these problems are solved, investigators must construct models of classification which conform best to the weight of theory and experience in this area. Given the current rather immature state of knowledge and technology in the field, one has cause to question why it is that the typology persists as the guiding structural model for psychopathology. As brought out at the conference, the model draws from two major sources of strength. One is that the weight of theory in the disciplines of personality and psychiatry seem to be heavily in support of the typological model. The major theories in these fields are typological theories. In psychiatry, for example, whether the orientation is descriptive, based on the ideas of Kraepelin, or whether it is psychodynamic, based on the concepts of Freud or Jung, the structural model underlying these theories is the same. These theories hold that qualitatively distinct types of personalities or mental disorders exist and are identifiable in terms of patterns of characteristics. It is true that theories based on other models exist, but it would be fair to say that current theory and clinical experience in this field are on the side of the typological model.

The second major source of strength for the model was that pointed to by Kety, Gerard, and Hamilton, if in somewhat different ways. It is simply that while recognizing the problems and dangers in the misapplication of classification, its role in the basic sciences is clear, i.e., "science regards classification as one of the primary prerequisites in the advancement of knowledge," "it is essential in structuring the search for more information and in . . . communicating what has been learned . . . [therefore] it must occur very early in the scientific process" (Kety). The kind of framework classification provides in the field of psychopathology facilitates other

types of research; e.g., research on the biochemical etiology of the schizophrenic psychosis.

Since the preferred theoretical model in psychopathology is the typological, the question posed to the statisticians was how best to implement theory in this respect. It was clear that definitive mathematical solutions to the mixture problem and to the definition of clusters are, as noted, not yet available. On the other hand, descriptive clustering procedures which have an extensive history in psychometrics are readily available and have been applied to several classes of data as described in the section on new systems. As clarified by Greenhouse, the fact that the clustering procedures are descriptive and inexact does not, in itself, rule out their utility. The techniques rather serve as the means whereby empirical, if untested, typologies are developed. The procedures are descriptive and based on approximate statistical techniques. The burden of establishing the validity of a typology based on a clustering approach lies with further experimental testing—the mathematics itself cannot unequivocally establish its existence. The fact that mathematical statistics has not as yet produced exact procedures for dealing with these problems is obviously not an argument for abandoning sound theory and experience in the field of psychopathology.

To determine, therefore, which statistical procedures are most appropriate, the purposes to which the research is directed have to be made very clear. Two major scientific but quite distinct research purposes have been identified; where the aim is to investigate and to validate the existence of natural groupings; where the aim is to solve a particular prediction problem. The statistical methodology, appropriate to each of these problems, can be very different, as was exemplified in the several research projects presented at the conference.

On this point we note, also, that discussants of the research projects presented in

the section entitled, "the development of new systems," had been asked to scrutinize very critically both the rationale and the techniques applied by these investigators. The discussants tended to be hypercritical so that developers of new systems found themselves attacked on the one hand for not dealing knowledgeably or in sufficient depth with clinical data (e.g., their failure to separate out the important and crucial symptoms from the less important, their tendency to think in terms of cross-sectional rather than in terms of longitudinal investigations, etc.), and on the technical side, for selecting statistical procedures in an arbitrary manner, for imputing too much validity to what are essentially nonprobabilistic, descriptive techniques. Yet these research investigators were also exposed at the conference to information that other, quite different, psychological methods had not been demonstrated to be very useful in this area (see the debate among Harris, Zubin, and Shakow, on the role of psychological tests in prognosis and classification), the fact already noted, that exact probabilistic techniques which were appropriate to the problem had not yet been developed, and to statements by the clinicians that, given the current state of methodology, we are still very much tied to subjective clinical observations as the primary data in this field. The work of the developers of new systems might, therefore, be viewed simply as attempts to codify and to quantify what is already known, in order that the field can move on to the development of other critical research in psychopathology.

Lehmann, for one, was, however, not satisfied with the manner in which clinical observations had been quantified by these investigators, and both he and Grinker argued that the primary observational data itself had to be improved. Yet they would not, in the light of the absence of etiologic information and the slow progress in the development of other psychological methods in this area, abandon the phenomenologic approach or the behavioral model

(Shakow). They were agreed, as were others at the conference, that we will have to continue to rely on the observer and the quantifying of manifest symptomatology and behavior as the base for other research in this area. In view of this, clinicians would have to be trained to be better observers, better naturalists, to, as Grinker put it, "resist theorizing and interpretation"—if improvement of clinical data was to be effected. The researchers would, in turn, have to understand that cross-sectional views of patients were inadequate, and intensive study of patients over time and consideration of phases of illness would have to eventually be integrated into any new typology.

The technical criticisms of the research papers, particularly those of Greenhouse, Lubin, Forgy, and Laska, were pointed and deserve careful study. They had been asked, as noted, to be hypercritical and reviewed the technical aspects of the research with great care. Lubin's dialogue and the massive complexity of the Gerard-Mattsson work, are very instructive in pointing up the fact that the current state of the field, both as regards content and technical development, demands a great deal of trial and error research as the only possible alternative to no research at all. Lubin's discussion covers many of the major unsolved problems in applying technique to theory in this area, and the Gerard-Mattsson program demonstrates the methodologic problems encountered when one tries to approach classification by going far beyond the cross-sectional data on behavior, exemplified by Lorr and Overall, to include all relevant measures of human functioning.

The situation cannot be nearly as bad as several of the critics would have it, but the research is obviously only in the early stages of its development. The fact that the nature of clinical observations, its recording, and its quantification are not yet at the level we would like them, does not change the fact that much progress has

already been accomplished in this area. The quantification of behavior and overt symptomatology has already led to significant research in etiology, in prediction, and in the evaluation of treatment. Given the current state of development, however, new typologies based on this data must be continually subjected to experimental tests of validity and refinement. Hamilton, Zubin, Katz, and Cole went into some detail on the problem of how validation of a typology might be accomplished and what the particular criteria might be. Noted in particular were the distinctions between consensual and predictive validity (Katz). The debate as to whether the criteria should be pragmatic (predictive) or theoretical recurs again, cautioning us further on how important it is to keep the aims of the particular research, the purposes of the classification project, in very sharp focus.

In closing, we note again that the conference goals were broad and it was not expected that final solutions of the major problems in this area would be accomplished. The orientations of psychiatrists, psychologists, and statisticians to the problem were known to be quite sharply in contrast and we felt it was sufficient, at this point in time, to expose them to each others' ways of looking at and dealing with a problem which is central to the development of the science of psychopathology and to the practice of psychiatry.

For the science to progress, the major issues in this field had to be articulated more clearly and the tendency for the disciplines of psychiatry and psychology to be out of touch with each others' perception of these issues represents a definite obstacle to such progress.

As a sourcebook the current volume should make quite clear how the various disciplines view the purpose of classification, what each discipline is trying to accomplish through the development of a system, how representatives of other related scientific disciplines; e.g., biology, sociology, view its role, what kinds of methods are being used to measure the relevant or significant phenomena—and the status of statistical and psychological methodology for dealing with the very complex technical problems which confront the investigators who are actually engaged in developing new systems. The conference ended with an overview and a consideration of the criteria with which new systems may be scientifically validated by representative members of the several disciplines.

In summary, then, the conference proceedings provide a base for assessing the current state of the science of classification in psychopathology and, hopefully, will provide a framework of theoretical and technical considerations against which future developments in the field can be compared and evaluated.

Appendix I

AMERICAN PSYCHIATRIC ASSOCIATION-NATIONAL INSTITUTE OF MENTAL HEALTH CONFERENCE PLANNING COMMITTEE

WALTER E. BARTON, M.D.
American Psychiatric Association
Washington, D.C.

Medical Director
American Psychiatric Association

HENRY BRILL, M.D.
Pilgrim State Hospital
West Brentwood, N.Y.

Chairman, Committee on Nomenclature and Statistics
American Psychiatric Association (1960-65)

JONATHAN O. COLE, M.D.
National Institute of Mental Health
Chevy Chase, Md.

Chief, Psychopharmacology Research Branch
National Institute of Mental Health;
Committee on Research
American Psychiatric Association

LEON EISENBERG, M.D.
Johns Hopkins University
Baltimore, Md.

Committee on Nomenclature and Statistics
American Psychiatric Association

LEON EPSTEIN, M.D.
Langley Porter Clinic
San Francisco, Calif.

Chairman, Committee on Research
American Psychiatric Association

BERNARD GLUECK, JR., M.D.
Institute of Living
Hartford, Conn.

Committee on Research
American Psychiatric Association

ERNEST GRUENBERG, M.D.
Columbia University
New York, N.Y.

Committee on Nomenclature and Statistics
American Psychiatric Association

SAMUEL W. GREENHOUSE, Ph. D.
National Institute of Mental Health
Chevy Chase, Md.

Assistant Chief, Office of Biometry
National Institute of Mental Health

MARTIN M. KATZ, Ph. D.
National Institute of Mental Health
Chevy Chase, Md.

Chief, Special Studies Section
Psychopharmacology Research Branch
National Institute of Mental Health

SAMUEL B. LYERLY, Ph. D.
Washington, D.C.

Research Associate
Bureau of Social Science Research, Inc.

Appendix II—Program

CONFERENCE ON THE ROLE AND METHODOLOGY OF CLASSIFICATION IN PSYCHIATRY AND PSYCHOPATHOLOGY

Washington Hilton Hotel, Washington, D.C., November 19–21, 1965

Under the Auspices of

THE AMERICAN PSYCHIATRIC ASSOCIATION AND
THE PSYCHOPHARMACOLOGY RESEARCH BRANCH OF
THE NATIONAL INSTITUTE OF MENTAL HEALTH

November 19

- 9:00 a.m. **WELCOMING REMARKS**
 Walter E. Barton, M.D.
 Raymond Feldman, M.D.
- 9:10 a.m. **THE RATIONALE AND GENERAL STRUCTURE OF THE CONFERENCE**
 Martin M. Katz, Ph. D.
 Jonathan O. Cole, M.D.
- 9:25 a.m. **1. THE HISTORY OF CLASSIFICATION IN THE MEDICAL SCIENCES**
 Oswei Temkin, M.D.

Section I

THE VARYING ROLES OF CLASSIFICATION IN PSYCHIATRY AND PSYCHOPATHOLOGY

Leon Epstein, M.D., Chairman

A. The Role of Classification in Psychiatric Practice

- 10:30 a.m. **2. THE ROLE OF CLASSIFICATION IN HOSPITAL PSYCHIATRY**
 Henry Brill, M.D.
- 10:40 a.m. **3. THE ROLE OF THE CLASSIFICATION SYSTEM IN OUTPATIENT PSYCHIATRY**
 Elmer A. Gardner, M.D.
- 10:50 a.m. **4. THE ROLE OF CLASSIFICATION IN PSYCHOANALYTIC PRACTICE**
 Harry Weinstock, M.D.
- 11:00 a.m. **OPEN DISCUSSION FOLLOWED BY ADJOURNMENT OF MORNING SESSION**

B. The Role of Classification in Research in Psychopathology

- 1:15 p.m. **5A. EPIDEMIOLOGY AND MEDICAL CARE STATISTICS**
 Ernest M. Gruenberg, M.D.
- 1:25 p.m. **5B. CLASSIFICATION OF MENTAL DISORDERS FOR EPIDEMIOLOGIC AND MEDICAL CARE PURPOSES—CURRENT STATUS, PROBLEMS AND NEEDS**
 Morton Kramer, Sc. D.
- 1:35 p.m. **6. THE ROLE OF CLASSIFICATION IN THE DEVELOPMENT OF THE SCIENCE OF PSYCHOPATHOLOGY WITH PARTICULAR REFERENCE TO RESEARCH**
 David Shakow, Ph. D.
- 1:45 p.m. **7. CLASSIFICATION IN RESEARCH ON THE PREDICTION OF RESPONSE TO SPECIFIC TREATMENTS IN PSYCHIATRY**
 Jonathan O. Cole, M.D.
- 1:55 p.m. **OPEN DISCUSSION**

C. *The Role and Problems of Classification in Related Disciplines and Their Relevance to Problems in Psychiatry and Psychopathology*

- 3:15 p.m. 8. THE ROLE OF CLASSIFICATION IN PERSONALITY THEORY
George A. Kelly, Ph. D.
- 3:30 p.m. 9. ABNORMAL PSYCHOLOGY
Kenneth Hammond, Ph. D.
- 3:45 p.m. 10. SOME IMPLICATIONS OF CLASSIFICATION IN SOCIOLOGY FOR PROBLEMS OF CLASSIFICATION IN PSYCHIATRY AND PSYCHOPATHOLOGY
John A. Clausen, Ph. D.
- 4:00 p.m. 11. DISCUSSION OF PAPERS by BRILL, WEINSTOCK, GLESER, OVERALL and HOLLISTER, STEIN and NEULINGER IN THE LIGHT OF RESEARCH IN HUMAN DEVELOPMENT
Lois B. Murphy, Ph. D.
- 4:15 p.m. 12. PROBLEMS IN PSYCHIATRIC NOSOLOGY FROM THE VIEWPOINT OF THE BIOLOGICAL SCIENCES
Seymour S. Kety, M.D.
- 4:45 p.m. 13. NO PRESENTATION
OPEN DISCUSSION FOLLOWED BY ADJOURNMENT FOR DAY

November 20

Section II

ISSUES IN THE METHODOLOGY AND STATISTICS OF CLASSIFICATION

Samuel W. Greenhouse, Ph. D., Chairman

- 9:00 a.m. 14. ON THE MEANING OF DISCRIMINATION, CLASSIFICATION, MIXTURE, AND CLUSTERING IN STATISTICS
Samuel W. Greenhouse, Ph. D.
- 9:25 a.m. 15. QUANTIFYING SIMILARITY BETWEEN PEOPLE
Goldine C. Gleser, Ph. D.
- 9:50 a.m. 16. MULDIMENSIONAL REPRESENTATION OF SIMILARITY STRUCTURES
Warren S. Torgerson, Ph. D.
- 10:30 a.m. 17. NO PRESENTATION
- 10:30 a.m. 18. A SURVEY OF SOME EMPIRICAL CLUSTERING PROCEDURES
Samuel B. Lyerly, Ph. D.
- 10:55 a.m. 19. DISCRIMINATION AMONG GROUPS AND ASSIGNING NEW INDIVIDUALS
C. Radhakrishna Rao, Sc. D.
- 11:20 a.m. OPEN DISCUSSION FOLLOWED BY ADJOURNMENT OF MORNING SESSION

Section III

THE CRITICAL REVIEW AND DISCUSSION OF THE METHODOLOGY INVOLVED IN A SERIES OF ATTEMPTS TO DEVELOP NEW TYPOLOGIES

George A. Kelly, Ph. D., Chairman

A. Descriptive and Phenomenological Approaches

- 2:00 p.m. 20. THE PHENOMENA OF DEPRESSIONS
Roy R. Grinker, Sr., M.D., and Jum C. Nunnally, Ph. D.
- 2:10 p.m. 21. A TYPOLOGY FOR FUNCTIONAL PSYCHOTICS
Maurice Lorr, Ph. D.
- 2:20 p.m. 22. STUDIES OF QUANTITATIVE APPROACHES TO PSYCHIATRIC CLASSIFICATION
John E. Overall, Ph. D., and Leo E. Hollister, M.D.
- 2:30 p.m. 23. A PHENOMENOLOGICAL TYPOLOGY OF SCHIZOPHRENIA
Martin M. Katz, Ph. D.
- 2:40 p.m. DISCUSSION OF PAPERS BY: GRINKER and NUNNALLY, OVERALL and HOLLISTER, KATZ
Ardie Lubin, Ph. D.
Samuel Greenhouse, Ph. D.
Heinz E. Lehmann, M.D.
- 3:45 p.m. RESPONSES BY GRINKER, NUNNALLY, LORR, OVERALL, and KATZ
- 4:15 p.m. OPEN DISCUSSION FOLLOWED BY ADJOURNMENT OF AFTERNOON SESSION

B. Systems Based on Patterns of Psychological Test Performance

- 8:00 p.m. 24. BIOMETRIC ASSESSMENT OF MENTAL PATIENTS
Joseph Zubin, Ph. D.
- 8:10 p.m. 25. PATTERNS OF INTELLECTUAL FUNCTIONING AND THEIR IMPLICATIONS FOR THE DYNAMICS
OF BEHAVIOR
David R. Saunders, Ph. D., and John W. Gittinger, M.S.
- 8:20 p.m. 26. A TYPOLOGY OF SELF-DESCRIPTIONS
Morris I. Stein, Ph. D., and John Neulinger, Ph. D.
- 8:30 p.m. DISCUSSION OF PAPERS BY: ZUBIN, SAUNDERS and GITTINGER, STEIN and NEULINGER
Robert E. Harris, Ph. D.
Edward Forgy, Ph. D.
- 9:30 p.m. ADJOURNMENT FOR THE DAY

November 21

C. Systems Based on Premorbid History, Course of Illness or Reaction to Treatment

- 9:00 a.m. 27. PROCESS AND REACTIVE SCHIZOPHRENIA—SOME CONCEPTIONS AND ISSUES
Norman Garmezy, Ph. D.
- 9:10 a.m. 28. DIAGNOSIS AND PATTERN OF REACTION TO DRUG TREATMENT—CLINICALLY DERIVED
FORMULATIONS
Donald F. Klein, M.D.
- 9:20 a.m. 29. PATTERNS OF REACTION TO DRUG TREATMENTS DERIVED THROUGH MULTIVARIATE PROCE-
DURES
Dean J. Clyde, Ph. D.
- 9:30 a.m. DISCUSSION OF PAPERS BY GARMEZY, KLEIN, and CLYDE
Heinz Lehmann, M.D.
C. James Klett, Ph. D.

*D. Systems Which Attempt To Encompass Variables From Several of the Significant Dimensions
of Functioning in Mental Illness*

- 10:45 a.m. 30. TYPOLOGY OF SCHIZOPHRENIA BASED ON MULTIDISCIPLINARY OBSERVATIONAL VECTORS
Nils B. Mattsson, LL.M., and Ralph W. Gerard, M.D.
- 10:55 a.m. 31. NEUROPHYSIOLOGICAL RESPONSE STRATEGIES IN THE CLASSIFICATION OF MENTAL ILLNESS
Max Fink, M.D.
- 11:00 a.m. DISCUSSION OF PAPERS BY: MATTSSON and GERARD, FINK
Eugene Laska, Ph. D.
Max Hamilton, M.D.
- 12:30 p.m. ADJOURNMENT OF MORNING SESSION

Section IV

THE VALIDITY AND UTILITY OF CLASSIFICATION SYSTEMS

Benjamin Pasamanick, M.D., Chairman

- 2:00 p.m. 32. DISCUSSION ON CONFERENCE
Max Hamilton, M.D.
- 2:30 p.m. 33. A TYPOLOGY OF MODELS
Max A. Woodbury, Ph. D.
- 3:00 p.m. DISCUSSION
Jonathan O. Cole, M.D.
Martin M. Katz, Ph. D.
Joseph Zubin, Ph. D.
- 3:30 p.m. OPEN DISCUSSION
- 4:30 p.m. CONFERENCE ADJOURNED

Appendix III

ALPHABETICAL ROSTER ¹

CONFERENCE ON THE ROLE AND METHODOLOGY OF CLASSIFICATION IN PSYCHIATRY AND PSYCHOPATHOLOGY

Washington, D.C., November 1965

Under the Auspices of

THE AMERICAN PSYCHIATRIC ASSOCIATION AND THE PSYCHOPHARMACOLOGY RESEARCH BRANCH OF THE NATIONAL INSTITUTE OF MENTAL HEALTH

WALTER E. BARTON, M.D., Medical Director, American Psychiatric Association, Washington, D.C.

HENRY BRILL, M.D., Director, Pilgrim State Hospital, West Brentwood, Long Island, N.Y.

JOHN A. CLAUSEN, Ph. D., Professor of Sociology, Institute of Human Development, University of California, Berkeley

DEAN J. CLYDE, Ph. D., Director, Biometric Laboratory, University of Miami, Coral Gables, Fla.

JONATHAN O. COLE, M.D., Chief, Psychopharmacology Research Branch, National Institute of Mental Health, Chevy Chase, Md.

LEON EPSTEIN, M.D., Associate Medical Director, Langley Porter Neuro-Psychiatric Institute, University of California Medical Center, San Francisco

RAYMOND FELDMAN, M.D., Deputy Director, National Institute of Mental Health, Bethesda, Md.; subsequently, Director, Mental Health Programs, Western Interstate Commission for Higher Education, Boulder, Colo.

MAX FINK, M.D., Professor of Psychiatry, New York Medical College, New York, N.Y.

EDWARD FORGY, Ph. D., Research Specialist, Statistics, Department of Psychiatry, University of California Medical Center, Los Angeles

ELMER A. GARDNER, M.D., Associate Professor and Director, Community Mental Health Center, Department of Psychiatry, Temple University Health Science Center, Philadelphia, Pa. Formerly: Associate Professor and Director, Division of Preventive Psychiatry, Department of Psychiatry, University of Rochester School of Medicine and Dentistry, Rochester, N.Y.

NORMAN GARMEZY, Ph. D., Director, Center for Personality Research, Department of Psychology, University of Minnesota, Minneapolis

RALPH W. GERARD, M.D., Dean of the Graduate Division, University of California, Irvine

JOHN W. GITTINGER, M.S., Psychological Assessment Associates, Washington, D.C.

GOLDINE C. GLENER, Ph. D., Professor of Psychology, Department of Psychiatry, University of Cincinnati College of Medicine, Ohio

SAMUEL W. GREENHOUSE, Ph. D., Chief, Theoretical Statistics and Mathematics Section, Office of Biometry, National Institute of Mental Health, Chevy Chase, Md.

ROY R. GRINKER, Sr., M.D., Director, Institute for Psychosomatic and Psychiatric Research and Training, Michael Reese Hospital and Medical Center, Chicago, Ill.

ERNEST M. GRUENBERG, M.D., Dr. P.H., Department of Psychiatry, Columbia University College of Physicians and Surgeons, New York, N.Y.

MAX HAMILTON, M.D., Nuffield Professor of Psychiatry, University of Leeds School of Medicine, England

KENNETH R. HAMMOND, Ph. D., Institute of Behavioral Science, University of Colorado, Boulder

ROBERT E. HARRIS, Ph. D., Professor of Medical Psychology, University of California School of Medicine, San Francisco

LEO E. HOLLISTER, M.D., Veterans Administration Hospital, Palo Alto, Calif.

MARTIN M. KATZ, Ph. D., Chief, Special Studies Section, Psychopharmacology Research Branch, National Institute of Mental Health, Chevy Chase, Md.

GEORGE A. KELLY, Ph. D., Professor of Psychology, Department of Psychology, Brandeis University, Waltham, Mass.

¹ Lists of Guest Participants and Auditors follow.

SEYMOUR S. KETY, M.D., Chief, Laboratory of Clinical Science, National Institute of Mental Health, Bethesda, Md.

DONALD F. KLEIN, M.D., Director of Research, Hillside Hospital, Glen Oaks, N.Y.

C. JAMES KLETT, Ph. D., Chief, Central Neuropsychiatric Research Laboratory, Veterans Administration Hospital, Perry Point, Md.

MORTON KRAMER, Sc. D., Chief, Biometry Branch, Office of Program Planning and Evaluation, National Institute of Mental Health, Public Health Service, Department of Health, Education, and Welfare, Chevy Chase, Md.

EUGENE LASKA, Ph. D., Director, Computer Center, Rockland State Hospital, Orangeburg, N.Y.

HEINZ E. LEHMANN, M.D., Clinical Director, Douglas Hospital, Verdun, and Professor of Psychiatry, McGill University, Montreal, Canada

MAURICE LORR, Ph. D., Professor of Psychology and Psychiatry, Catholic University, Washington, D.C. Formerly: Chief, Outpatient Psychiatric Research Laboratory, Veterans Administration, Washington, D.C.

ARDIE LUBIN, Ph. D., Research Psychologist, U.S. Navy Medical Neuropsychiatric Research Unit, San Diego Naval Hospital, California

SAMUEL B. LYERLY, Ph. D., Psychological Assessment Associates and Bureau of Social Science Research Inc., Washington, D.C.

NILS MATSSON, LL.M., Biometric Laboratory, George Washington University, Washington, D.C. Now with Advanced Systems Development Division, IBM Corp.

LOIS B. MURPHY, Ph. D., Research Psychologist, The Menninger Foundation, Topeka, Kans.

JOHN NEULINGER, Ph. D., Research Assistant Professor of Psychology, Research Center for Human Relations, Department of Psychology, New York University, New York City

JUM C. NUNNALLY, Ph. D., Professor of Psychology, Vanderbilt University, Nashville, Tenn.

JOHN E. OVERALL, Ph. D., Director, Research Computer Center, University of Texas, Medical Branch, Galveston

BENJAMIN PASAMANICK, M.D., Division of Research Services, Illinois State Department of Mental Health, Chicago

C. RADHAKRISHNA RAO, Sc. D., F.R.S., Director, Research and Training School, Indian Statistical Institute, Calcutta, India

DAVID R. SAUNDERS, Ph. D., Professor of Psychology, University of Colorado, Boulder

DAVID SHAKOW, Ph. D., Chief, Laboratory of Psychology, National Institute of Mental Health, Bethesda, Md.

MORRIS I. STEIN, Ph. D., Professor of Psychology, Research Center for Human Relations, New York University Graduate School of Arts and Science, New York, N.Y.

OWSEI TEMKIN, M.D., William H. Welch Professor of the History of Medicine, The Johns Hopkins University Institute of the History of Medicine, Baltimore, Md.

WARREN S. TORGERSON, Ph. D., Department of Psychology, Johns Hopkins University, Baltimore, Md.

HARRY I. WEINSTOCK, M.D., Chief, Department of Psychiatry, Mount Sinai Hospital Services, City Hospital Center at Elmhurst, Elmhurst, N.Y.

MAX A. WOODBURY, Ph. D., Professor of Biomathematics, Duke University Medical School, Durham, N.C.

JOSEPH ZUBIN, Ph. D., Chief of Psychiatric Research (Biometrics), New York State Department of Mental Hygiene, New York, N.Y.

GUEST PARTICIPANTS

ANITA BAHN, Sc. D., Chief, Outpatient Studies, Office of Biometry, National Institute of Mental Health, Bethesda, Md.

GEOFFREY H. BALL, Ph. D., Senior Research Engineer, Radio Systems Laboratory, Stanford Research Institute, Menlo Park, Calif.

ARTHUR BRAYFIELD, Ph. D., Executive Officer, American Psychological Association, Washington, D.C.

JOSEPH FLEISS, M.S., Associate Biostatistician, Biometrics Research, New York State Department of Mental Hygiene, New York, N.Y.

SOL L. GARFIELD, Ph. D., Director, Clinical Psychology Program, Teachers College, Columbia University, New York, N.Y.

BERNARD C. GLUECK, JR., M.D., Director of Research, The Institute of Living, Hartford, Conn.

BURGESS L. GORDON, M.D., Director, Office of Current Medical Terminology, American Medical Association, Chicago, Ill.

PAUL J. HOFFMAN, Ph. D., Director, Oregon Research Institute, Eugene

B. R. HUTCHESON, M.D., Director, Division of Mental Hygiene, Massachusetts State Department of Mental Health, Boston

RICHARD L. JENKINS, M.D., Chief, Child Psychiatry, State University of Iowa College of Medicine, Iowa City

GERALD L. KLERMAN, M.D., Associate Professor of Psychiatry, Yale University School of Medicine, New Haven, Conn.

LAWRENCE S. KUBIE, M.D., Baltimore, Md.

HENRI A. LOWERY, M.S., Psychologist, Psychopharmacology Research Branch, National Institute of Mental Health, Chevy Chase, Md.

LOUIS L. MCQUITT, Ph. D., President, Division of Evaluation and Measurement, American Psychological Association and Dean, College of Social Science, Michigan State University, East Lansing

BERNARD RIMLAND, Ph. D., Director, Personnel Measurement, Research Department, U.S. Naval Personnel Research Activity, San Diego, Calif.

MYRON SANDIFER, Jr., M.D., Biometrics Research, New York State Department of Mental Hygiene, New York, N.Y.

ROBERT L. SPITZER, M.D., Senior Research Psychiatrist, New York State Psychiatric Institute, New York, N.Y.

HAROLD M. VISOTSKY, M.D., Director, Illinois State Department of Mental Health, Chicago

BEATRICE WERBLE, Ph. D., Michael Reese Hospital and Medical Center, Chicago, Ill.

AUDITORS

ROBERT ANDERSON, M.D., Psychopharmacology Research Branch, National Institute of Mental Health, Chevy Chase, Md.

MITCHELL BALTER, Ph. D., Psychopharmacology Research Branch, National Institute of Mental Health, Chevy Chase, Md.

THEODORE H. BLAU, Ph. D., President, Division of Consulting Psychology, American Psychological Association, Tampa, Fla.

ROLAND BONATO, Ph. D., Director, Biometric Laboratory, George Washington University, Washington, D.C.

SPENCER M. FREE, Jr., Ph. D., Head, Statistics Section, Smith Kline & French Laboratories, Philadelphia, Pa.

SOLOMON GOLDBERG, Ph. D., Psychopharmacology Research Branch, National Institute of Mental Health, Chevy Chase, Md.

DONALD W. HAMMERSLEY, M.D., Chief, Professional Services, American Psychiatric Association, Washington, D.C.

GEORGE HENINGER, M.D., Office of the Director, National Institute of Mental Health, Chevy Chase, Md.

EINAR KRINGLEN, M.D., Adult Psychiatry Branch, National Institute of Mental Health, Chevy Chase, Md.

JEROME LEVINE, M.D., Psychopharmacology Research Branch, National Institute of Mental Health, Chevy Chase, Md.

RONALD LIPMANN, Ph. D., Psychopharmacology Research Branch, National Institute of Mental Health, Chevy Chase, Md.

JOHN J. McMILLAN, Ph. D., Administrative Officer for State and Professional Affairs, American Psychological Association, Washington, D.C.

RUSSELL MONROE, M.D., The Psychiatric Institute, Baltimore, Md.

JAMES PRESCOTT, Ph. D., Office of Naval Research, Washington, D.C.

ROBERT PRIEN, Ph. D., Biometric Laboratory, George Washington University, Washington, D.C.

ALLEN RASKIN, Ph. D., Psychopharmacology Research Branch, National Institute of Mental Health, Chevy Chase, Md.

CHARLES E. RICE, Ph. D., Biometric Laboratory, George Washington University, Washington, D.C.

THOMAS SONN, M.D., Psychopharmacology Research Branch, National Institute of Mental Health, Chevy Chase, Md.

HUSSEIN TUMA, Ph. D., Division of Research Grants, National Institute of Mental Health, Chevy Chase, Md.

IRENE WASKOW, Ph. D., Psychopharmacology Research Branch, National Institute of Mental Health, Chevy Chase, Md.

PAUL T. WILSON, M.D., Principal Investigator, Information Processing Project, American Psychiatric Association, Washington, D.C.

Index

Numerals in **BOLDFACE** indicate major entries.

Numerals in *italics* indicate entries in figures or illustrations.

- Abnormal psychology, **164**
- Acetophenazine, 289,295,297,487,488,489
- "Activation" mechanisms, 474,480,481,482
- Addictions, 66. *See also* Alcohol; Barbiturates
- Adler, Alfred, 403n
- Administration, 3,85,164,166. *See also* Hospital administration
- Admission rates, 104
 - by diagnosis, 105,106
 - See also* under name of disease classification
- Adrenergic mechanisms, 535,537
- Affective psychosis, 41,46,48,106,106
 - admission rates of, 105,106
 - in old age, 109
- Aged, 73,80,101,103,108,109,557. *See also* Involutional psychotic reactions
- Agents (in epidemiology), 77
 - classification of, 78
- Alcohol and alcoholism, 41,48,103,104-105,128n,410n,-425,477,512
 - and outpatient care, 47
- Alexander, Franz, 132
- Allport, Gordon W., 122,136
- American Heart Association nomenclature, 4,9,70,72,-73,74
- American Hospital Association, 112,114
- American Medical Association, 150
- American Medico-Psychological Association, 36
- American Orthopsychiatric Association, 73,148
- American Psychiatric Association. *See* APA
- American Psychoanalytic Association, 134
- American Psychological Association, 134
- American Public Health Association, 79
- Amitriptyline, 295,297
- Amobarbital, 536
- Analytic approach to clustering, 222
- Anderson, Robert, 575
- Androgyny score, 526n
- Anxiety and anxiety reactions, 66,332
 - and behavior, 159,160
 - and chlordiazepoxide, 143
 - chronic, 477
 - episodic, 469
 - severe, 91
 - and test results, 7
 - and thioridazine, 324
- Anxious intropunitiveness, 263,264
- APA, 4,37n,60,71-72,100,129,134
 - classification, 34,39,73,90,93,131,148,350
 - diagnostic system, 37,67,184
 - Manual, 30,34,63,99,112,276
 - Nomenclature, 4,9,23,40,49,101,346
 - See also* Standard APA Nomenclature
 - nosology, 33
 - and psychoanalysis, 63,64
- Arbitrary decisions, 541,542
- Aretaeus, 13
- Aristotle, 13,95
- Armed Forces nomenclature, 37,51. *See also* U.S. Army classification
- Arteriosclerotic psychoses, admission rates of, 107
- Asch, S.E., 399
- Assessment centers, 559
- Atropine, 537
- Autism. *See* Children, autistic
- "Autism in Children With Mental Deficiency" (Bender), 176
- B* coefficient, 221,517
- Bahadur, R., 231
- Bahn, Anita, 73,148,149,245,574
- Balinsky, B., 378
- Ball, Geoffrey H., 2,244,245,322,349,574
- Balter, Mitchell, 575
- Barbiturates, 477,535,536
- Barrabee, E., 113
- Barrabee, P., 113
- Barrabee and Finesinger Social Adjustment Scale, 113
- Barron, F., 435
- Barton, Walter E., 82,503,569,570,573
- "Basic Psychiatric Literature . . . of Residency Training Programs," 57n.
- "Basis for the Improvement of Personality Tests . . . (Rosenzweig), 130
- Bateson, G., 357,531
- Bayesian technique, 199,277,282,328
- Bayley, Nancy, 179
- BD. *See* Block Design Subtest

- Beck, Sam, 151,275,531
 Becker, W.C., 428,432,434,435,439,449,461
 Behavior disorders, 145. *See also* under Children
 Behavior Profile Rating Scale, 5
 Behavior ratings, 249,250,304,332
 Behavioral sciences, 134
Belle indifference, 73,407
 Bender, Laurretta, 174,176,177,182,543
 Bender, J.L., 286
 Benjamin Proverb Test, 405,434,439,442,443
 Bennett, J.L., 286
 Benton, A.L., 369
 Benzquinamide, 289
 Berger, Frances, 398
 Berlin Psychoanalytic Institute, 62
 Bernard, Claude, 17
 Bichat, Xavier, 15
 Biochemistry, 556
 Biological sciences, 8,117,133,190,567
 "Biometric Approach to Prognosis in Schizophrenia, A" (Hoch and Zubin), 505
 Biometric Research Group, 362-363
 Biometrics, 332,342,353
 Biophobia, 434
 Biostatistics, 99,102,103,353
 Birren, J.E., 378
 Blake, R.R., 399n
 Blau, Theodore H., 575
 Bleuler, Eugen, 29,88,91,92,96,123,138,249,419,422,-440,451,507
 Bloch's law, 371
 Block Design Subtest, 382
 Body types, 202
 Boisen, Anton, 119
 Boissier de Sauvages, F., 13,14
 Bonato, Roland, 575
 Borderline states, 66,260,470
 and psychoanalysis, 67,70
 Bordin, E.S., 132
 Boston Psychopathic Hospital, 119,120,121,122,131,-151
Bouffée delirante, 104
 Bourdon Dot Cancellation, 526-527
 Bowman, Karl M., 119,121
 BPRS. *See* Brief Psychiatric Rating Scale
 Bradford, W., 432n
 Brain damage syndrome, 93,104,104,369
 acute, 42,45,51
 children's. *See* under Children
 chronic, 41,45,48,50,51
 Brain disorders, 42,91
 Brain function approach, 355
 Braren, Margery, 360
 Brayfield, Arthur, 574
 Brehm, J.W., 399n
 Brief Psychiatric Rating Scale, 277,278,279,280,281,-286,287,290,291,292,295,296,297,338
 Brill, Henry, 2,3,20,23,25,29,52,72,73,74,99,148,149,-173,184,483,569,570,573
 Brooklyn State Hospital, 363
 Brune, G.G., 474
 Brunswik, Egon, 167
 Bunsen-Roscoe law, 371
 Burdock, E.L., 262,302,359,363,364,366,415
 Burdock-Ward Behavior Scale, 504
 Burnaby, T.P., 235
 Burstein, A., 427,430
 Burt, C.L., 509n
 Byrne, D., 344
 Cabot, Richard C., 119,120
 Caelius Aurelianus, 13,21
 Caffey, Gene, 286
 California F scale, 537n
 California Psychological Inventory, 404
 Calvi, Gabriele, 399n
 Cameron, N., 275,441n
 Campbell, MacFie, 119
 Cancro, Robert, 419n
 Canonical analysis, 9,283,327,482,487
 Carrigan, P.M., 382
 Carroll, Lewis, 140,141
 Case registers, 105. *See also* under name
 Categorization, 117,157,158,159,164
 Dewey system of, 118
 Cattell, R.B., 207,224,262,349,379,509n
 Cavalli-Sforza, L.L., 226
 Central fact-gathering committee (APA), 63,71-72
 "Centration", 449
 Chamberlin, T.C., 135
 Character disorders and neuroses, 66,67
 and emotional instability, 471,472,473
 and psychoanalysis, 62
 Chapman, L.J., 427,428,429,430,431,434
 Chassan, Jack B., 62
 Chicago Psychoanalytic Institute, 62
 Child and children, autistic, 175,176,177,243
 behavior disorders of, 81,95,149
 borderline psychotic, 177
 brain-damaged, 81,176,177
 convulsive disorders of, 81
 mental retardation of, 81
 pattern of growth of, 175
 schizophrenic, 176,177,357
 therapy, 178
 Children's Behavior Inventory, 363,366
 Chloralose, 535
 Chlordiazepoxide, 143,477
 Chlorpromazine, 191,314,329,348,468,488,495,496,497
 Cholinergic drugs, 535,537
 Choynowski, Mieczyslaw, 399n
 Chronic brain syndrome. *See* Brain damage syndrome, chronic
 Church, J.C., 436

- Classification, of cases, 77,83,94,97
 of conditions, 83
 of environments, 78,97
 European and international, 35,93,112,150
 of groups, 239
 history of, 11,26
 in medicine, 551
 methodology of, 195
 multidimensional, 5,51,74
 vs. nomenclature, 148
 premature, 117
 problem, 132,134,508,509,410
 factor-analytic approach to. *See* Factor analysis
 and prognosis, 175
 psychosocial, 148
 purpose of, 27,102,563
 vs. recording, 74,75
 and statistics, 195,197,198
 systems, noncategorical, 162
 recent efforts at, 131,247
 and treatment, 174
See also under name
- Clausen, John A., 5,168,188,189,571,573
- Clinical interview, structured. *See* Structured Clinical Interview
- Clinical psychology, 1,197,201,328,553,554,563
- Clustering and clusters, 160,195,197,199,200,201,204,-213,220,227,264,411
 analysis, 6,219,327,334,346,347,407
 mathematical model underlying, 558
 modal, 280,287
 of order one, 226
 procedures and techniques, 222,224,241,242,244,-508,511,517,518,565
 of traits, 200
- Clyde, Dean J., 9,245,313n,483,490,491,493,494,495,-500,502,505,560,572,573
- CMT. *See* "Current Medical Terminology"
- Coan, R. W., 132
- Cochran, W.G., 231
- Coefficient of belonging. *See* B coefficient
- Cohen, Henry, 136
- Cohen, J., 378,387,396n
- Cohn, Ferdinand, 16
- Cohnheim, Julius, 17
- Cohort analysis, 504
- Cole, Jonathan O., 1,3,9,143,169,225,279,282,300n,-326,363,491,504,556,563,567,569,570,572,573
- College of Neuropsychopharmacology, 501
- Commission on Professional and Hospital Activities (Ann Arbor, Mich.), 56n
- Committee on Psychosocial Diagnoses and Classification (Amer. Ortho. Association), 148
- Community diagnosis, 79
- Community mental health centers, 105,112
- Community Mental Health Centers Act, 113
- "Comparative Epidemiology in the Mental Disorders" (Hoch and Zubin), 505
- Compensation (PAS term), 380,381
- Component analysis, 541
- Comprehension test, 378,382
- "Comprehensive Psychiatric Programs: A Survey..." (APA and NIMH), 37n
- Compulsive personality, 66. *See also* under Children
- Computers, 2,40,147,155,163,174,186,187,188,189,241,-257,282,338,339,416,518,533,534
 in psychoanalytic study, 63,66,67
- Comrey, A.L., 407
- Conceptual disorientation, 263,264
- Conceptual tests, 359,442
- Condillac, E.B., 14
- Conditioning approach, 355,357
- Conference, 1,551,556,559,563
- Configuration, effect of coefficient on, 217
- Confusion and confusional states, 91
- Conger, J.J., 225,226,262
- Construction matrix, 155
- Continuous variables, 389
- Conversion reaction, 64,66,67
- Cooley, W.W., 510
- Creativity group and research, 392,399
- Cronbach-Gleser distance function, 200
- Cronbach, Lee J., 205,211,221,262,284, 286,305
- Crozier, W.J., 139n
- Cullen, William, 13,14,16,17
- Cultural stereotypes. *See* under Stereotypes
- Culture-dependent, culture-fair, culture-free techniques, 360,367,368,375,417
- "Current Medical Terminology", 150
- Cushing, Jean, 63
- Cyclothymic personality, 43,66,67
- Darwin, Charles, 5,21,191
- Davis, P.C., 378
- Day, D., 427,430
- Decision rules, 282
 See also Optimum decision rule
- Defects, 3,24,75,557
- Definitions, 94,95
- Degan, J.W., 131
- Delany, M.J., 235
- Démence précoce*, 88,91
- Dementia praecox, 88,95,120,372,552
- "Dementia Praecox-Schizophrenia Problem" (May), 120
- Deming, Julia, 119, 120
- Density functions, 230,231
- Depression and depressive disorders, 8,66,95,106,143,-145,249,297,298,299,332
 admission rate of, 105,106
 agitated, 470
 endogenous, 192
 and oxypertine, 324,476
 prevention and control of, 105
 and psychoanalysis, 249
 psychotic, 39
 reactive, 192

- Depression and depressive disorders—Continued
 research, 250
 retarded, 8,468
 and thioridazine, 324
- Descriptive approach, 249
- Desmethylinipramine, 476
- Developmental approach, 174,178,355,356
 in child psychiatry, 180
- Dewey, John, 151
- Dewey system of categorization, 118
- Diagnosis and diagnostic, 3,4,19,22,71,73,102,167,256,-
 509,553,556
 biometric approach to, 8,355,373,374,375
 dialog, 324
 distribution, 103,107
 and drug effect, 477,478,479
 groups, 66,145
 in hospital administration, 30
 mathematical models for, 257
 multiple, 98,148
 objective approach to, 374
 psychosocial, 149
 reliability and validity of, 244,558
 standardization of, 102,103
 types, 332
 wastebasket, 43,72
- “Diagnostic and Statistical Manual: Mental Dis-
 orders” (APA), 100
- “Diagnostic and Statistical Manual of Mental Dis-
 eases” (Raines), 36,43
- “Dictionary of Occupational Titles”, 454
- “Dictionary of Psychological Medicine” (Tuke),
 116
- Digit span symbol and test, 378,382,446
- Dimensional approach, 214,215
- Disability vs. level of functioning, 98
- Discernment theory, 508
- Discriminant function analysis, 6,195,197,198,235,-
 243,289,482,509,511,516,518,519,520,521,523,524,525,
 556,564
See also Discrimination; Multiple discriminant
 analysis
- Discrimination, 197,200,509,511
 of composite hypotheses, 237
 among groups, **229**
See also Discriminant function analysis; Mul-
 tiple discriminant analysis
- Disease and diseases, 3,24,75,201,557
 nomenclature, 100
 statistical classification of, 100,101
- Disease entities vs. behavior patterns, 557
- “Disease and Injury Code” (Dept. of Defense), 51
- Disorders, 75,95,552
 classification of, 94,97,98
- Disorientation, 263,264
- Dissimilarity, elimination of, 208
- Dissociative reaction, 64,66,67
- Distance and distance functions, 7,200,205,226,227,-
 239,324
See also under name
- Distribution, bivariate normal and clustered, 204
- Ditran, 537
- Drug and drugs, 191,286,466,477,482,535,559
 and behavior, 191
 —devised-typology-devotee, 324
 evaluation, 497,499
 and NIMH-PSC study, 307,313
 and psychopathology, 490
 —refractory patients and states, 475,481
 response, 312,314,314,**466**,467,**483**,484,485,486,-
 486,487,487
 irregularities in, 477
 prediction of effects of, 481,489
 types, 281,288
See also NIMN-PSC study; under name of drug
- Duffy, E., 474
- Duke Schizophrenia Project, 431,452,453,453,456
- Durham decision, 89
- Edwards, A.W.F., 226
- Edwards Personal Preference Schedule, 404
- Edwin Smith Papyrus, 11
- Ego Strength Scale, 435
- Ehrlich, Paul, 19
- Eisenberg, Leon, 569
- Eisenbud, Jules, 63
- Eissler, Kurt, 133
- Ekman, P., 405
- Electroencephalography, 533,535,536,537,438,552
- Elgin Prognostic Scale, 426,427,428,429,430,431,433,-
 434,435,437,439,447,450,456,**461**,494,495,499
- Elgin State Hospital, 502
- “Eliminating Superfluous Variables . . .” (Lubin
 and May), 326
- Emotional inadequacy, 73
 and unstable personality, 64
- Empedocles, 11
- Engle, Ralph, 36
- Entitation of the universe, 153,154
- Environments (in epidemiology), 77
 classification of, 78,79
- Epidemiology, 3,**76**,78,83,86,97,99,200,332,358,505,551,
 563
 classification for, 99
 and public health practice, 79
- Epinephrine, 535,537
- EPS. *See* Elgin Pragnostic Scale
- Epstein, Leon, 20,72,147,148,150,569,570,573
- Epstein, S., 441n
- Equalifinality in personality theory, 402
- Erikson, E., 181
- Esquirol, J.E.D., 18
- “Evaluation of Psychoanalytic Therapy”, (psycho-
 analytic committee), 63
- Escalona’s infancy scale, 543
- Ethchlorvinyl, 477

- Euclidean distance and space, 218,225
- Excitement, 263,264
vs. retardation syndrome, 495,496,497
- Expert Committee on Health Statistics (WHO), 102
- Eysenck, H. J., 131,132
- Factor analysis, 6,139,168,199,218,224,227,252,276,327,-
328,338,339,341,346,347,372,378,392,406,510,556
raw score, 225
vs. typological analysis, 557
- Factorial screening technique, 509
- Farina, A., 431,432,464
- Fast-wave (beta) activity, 536
- "Fat stages", 175
- "Feeble-mindedness: Absolute, Relative, and Apparent" (Kanner), 176
- Feldman, Raymond S., 367,368,570,573
- Fine, H.J., 449
- Finesinger, J.E., 113
- Fink, Max, 9,73,74,474,535,540,542,543,547,548,572,-
573
- First admissions, by age and sex, 108,110,111
comparisons, 107
rates, 104,104,109
See also under name of disease classification
- Fisher, J., 406
- Fleiss, Joseph, 74,240,241,245,262,375,574
- Flicker frequency, 438
- Flint, Austin, 16,23
- Forgy, Edward W., 226,403,410,415,416,540,547,566,-
572,573
- Frank, J., 419
- Frazier, Shervert H., 57n
- Fluphenazine, 314,317,317,329,340,488,495,496,497,498
- Fortran computer program, 518
- Fragmentation, 501
- Free, Spencer M. Jr., 575
- Freud, Anna, 127n
- Freud, Sigmund, 62,68,71,72,88,90,92,331,344,356,-
403n,421,505,559,565
- Fromm, Erich, 403n
- Functional psychoses, 104,134,261,346
admission rates for, 105
- Functioning, level of, 98
- Funkenstein test, 531,535,542,543
- Galen, 12,13,17,202
- Galvanic skin response, 535
- GAP, 259,260
- Gardner, Elmer A., 4,23,35,52,72,73,74,169,552,570,-
573
- Garfield, Sol L., 75,574
- Garmezy, Norman, 1,5,9,419,431n,432,434,440,448,-
464,491,494,499,502,503,505,531,572,573
- Gault, U., 378
- Geisser, S., 309
- Genain quadruplets, 445
- "Genera Morborum" and "Genera Plantarum" (Linnaeus), 191,201
- General paresis, 87,88,91,95,145,419
- "General Psychopathology" (Jaspers), 22
- Genetic approach and genetics, 6,355,357,443,445,-
543,556
- Gerard, Margaret, 138
- Gerard, Ralph W., 5,9,151,152,153,404,405,445,482,-
507,508,532,540,542,547,565,566,572,573
- Gesell Developmental Examination, 544
- Gestalt psychology, 136,301
- Gibson, W.A., 224
- Giovannoni, Jeanne M., 431n
- Gittinger, John S., 8,377,381,383,403,404,572,573
- Gleser, Goldine C., 1,3,173,183,201,205,210,211,213,-
221,240,241,244,262,284,286,305,408,410,411,571,573
- Glueck, Bernard C. Jr., 245,491,500,505,569,574
- Glutethimide, 477
- Goldberg, Solomon, 282,302,502,504,575
- Goldfarb, W., 378
- Goldman, D., 535
- Goldstein, Kurt, 301,417
- Goldstein-Scheerer Sorting Test, 441,442
- Gordon, Burgess L., 76,150,243,574
- Gorham, Donald R., 132,278,280,282,296
- Gottschaldt Hidden Figures Test, 537n
- Gough, H.G., 404
- Grandiose expansiveness, 263,264
- Greenacre, Phyllis, 180,182,189,543
- Greenhouse, Samuel W., 1,6,72,73,197,220,240,241,-
243,245,309n,322,328,342,346,348,540,544,545,546,565,-
566,569,571,573
- Grinker, Roy S. Sr., 8,151,184,186,187,249,252,258,269,-
322,323,326,335,336,341,345,346,348,350,410,482,566,-
571,573
- Group and groups, assignment into, 223
classification of, 239
of schizophrenias, 92,372,424,552
- Group for the Advancement of Psychiatry. *See* GAP
- Gruenberg, Ernest M., 23n,76,97,147,148,149,169,-
172,491,503,569,570,573
- Guertin, W. H., 131
- "Guide to the Control of Mental Disorders" (Amer. Public Health Assoc.), 79
- Guilford, J.P., 379
- Hagen, E.P., 378
- Hakerem, G., 363,370
- Halstead, W.C., 406
- Hamilton, Max, 5,9,22,131,242,322,324n,350,351,540,-
542,547,549,551,559,565,567,572,573
- Hammer, A.G., 367,368
- Hammer, Muriel, 367
- Hammersley, Donald W., 575
- Hammond, Kenneth, 5,164,571,573
- Hardesty, Anne S., 363,366
- Hargreaves, W.A., 405
- Harlow, H.F., 357
- Harman, H.H., 221,509n, 511
- Haronian, F.J., 382

- Harris, J.G. Jr., 448,448,449,449,450
Harris, Robert E., 403,404,407,415,417,566,572,573
Hartmann, Heinz, 181
Healy, M.J.R., 235
Hebb, D.O., 361
Hebephrenia, 127
Heerman, Emil F., 207
Heider, Grace, 543
Heincke, F., 262
Hellinger's distance, 239
Henderson, David, 123,450
Heninger, George, 575
Henle, J., 15
Henry, G.W., 419
Herron, W.G., 434
Higgins, J., 433,434
Hill, Lewis, 138
Hillside Hospital, 466,467,480
Himwich, H.E., 474,481
Hinsie, Leland E., 90
Hippocrates, 11,13,136,555
Histrionic patients, 476
Hoch, P.H., 505
Hoffman, Paul J., 574
Holistic approach, 136,301
Hollister, Leo E., 5,173,183,277,280,322,323,324,325,-
326,328,331,338,342,571,573
Holzinger, K.J., 221,517
Hook, Marion E., 226
Hopkins, C.E., 231
Horwitt, M.K., 443,445
Hoskins, R.G., 450
Hospital Adjusting Rating Scale, 363
Hospital administration, 30,174
Hospital psychiatry, 29
Hosts (in epidemiology), 77
Hotelling analysis, 327
Hover, G.L., 378
Humoller rate, 520
Humors (in Greek medicine), 11-12,555
Humphreys, L.G., 132
Hunt, Edna B., 132
Hunt, W.A., 132
"Hunting of the Snark, The" (Carroll), 140
Huston, P.E., 425,427
Hutcheson, B.R., 574
Huxley, Julian, 557
Hypertension, 89,98
Hyperthymic psychopath, 472
Hysteria and hysterical personality, 43,481
Hysteroschizophrenia, 476-477
ICD. *See* International Statistical Classification of
Diseases, Injuries and Causes of Death
"Idiot research", 533
Identification problem, 509,511
Illinois psychiatric case register, 59n
Illinois State Psychiatric Institute, 260
Imipramine, 8,143,288,289,143,295,297,324,467,470,-
473,476,478,537n
and retarded, 324
IMPS. *See* Inpatient Multidimensional Psychiatric
Scale
Impulsive behavior disorders, in psychoanalytic
practice, 70
Incidence rate, 76
"Index to World Health Organization Classifica-
tion", 90
"Indications for Psychoanalysis" (seminar), 70
Information loss, 203,211,410
Information net, 156
Information subtest, 378
Inkeles, A., 399n
Inpatient Multidimensional Psychiatric Scale, 262,-
270,308n,309,310,349,495,504
Insanity, classification of, 18
treatise on, 29
Institute of Living, The, 56n
Institutes of Medicine (18th century), 12
Institutional neuroses, 93
Intellectual functioning, 377
Intelligence tests. *See* I. Q.
Interactional disorders, 51
Internal environmental approach to classification,
355
"International Classification of Diseases", 101
"International Classification of Mental Disorders",
112,150
"International List of Causes of Disease" (World
Health Organization), 35
International Revision Conference (WHO), 100
International Statistical Classification of Diseases,
Injuries and Causes of Death (WHO), 100,101,-
148,149,150
"Interpersonal Relations in Childhood" (question-
naire), 400
Intersubstitutability, 165,167
Interviews. *See* Structured Clinical Interviews
Involutional psychotic reactions, admission rates of,
105,106
I. Q., 7,174,175,179,359,388,425
and age at hospitalization, 474
Irreducible uncertainty, 165
Itil, T., 535
Jackson, Hughlings, 17,18,21,22,189
Jaspers, Karl, 22,353,354,419,442
Jelliffe, Smith Ely, 123
Jellinek, Morton, 127,128,129,130
Jenkins Prognostic Scale, 435
Jenkins, Richard L., 4,149,150,574
John, S., 231
John, Vera, 417
Johnson, Merlin, 286
Jung, C.G., 301,320,321,344,403n,421,565
Kahn Test of Symbol Arrangement, 520
Kaiser, H.F., 511

- Kanner, Leo, 176,243
 Kantor, R.E., 425,426,427,437,438,439
 Kasanin, J., 121
 Kato, Yachiyo, 399n
 Katz Adjustment Scales, 366
 Katz, George, 286
 Katz Behavior Inventories, 113
 Katz, Martin M., 1,5,8,9,113,225,**300**,301,303,305,320,-
 322,323,324,325,326,327,328,329,339,340,342,344,345,-
 348,359,366,482,540,546,547,**563**,567,569,571,572,573
 Kaufman, M.R., 134
 Keeler, W.R., 177
 Keller, L., 225,226,262
 Kelly, D.H.W., 475
 Kelly, George A., **155**,162,322,344,540,543,571,573
 Kempf, E.J., 121,123,131
 Kendall, M.G., 199,326,508
 Kent, Grace, 119
 Kety, Seymour S., 5,152,**190**,343,438,443,565,571,574
 Kimbell, Isham, 286
 King, G., 438,439
 Klein, Donald F., 3,9,245,**466**,474,481,491,493,494,-
 500,501,502,504,553,572,574
 Klerman, Gerald L., 23,73,152,322,351,483,491,492,-
 504,574
 Klett, C. James, 261n,262,263,269,326n,491,493,500,-
 501,502,504,564,572,574
 Kluckhohn, C., 136,137
 Knight, Robert P., 62,63
 "Knowing is naming", 151,189
 Koch, Robert, 16,192
 Kodama, H., 378
 Korsakoff's syndrome, 87
 Kornreich, M., 408
 Kossack, C.F., 508,509
 Kraepelin, Emil, 18,29,88,91,92,120,123,138,187,249,-
 331,337,351,352,354,372,419,420,421,422,424,450,451,-
 473,503,507,552,565
 Kramer, Morton, 74,76,84,**99**,107,147,149,150,169,-
 244,376,540,545,546,570,574
 Kretschmer, E., 202,353
 Kringlen, Einar, 575
 Kruskal, J.B., 214n
 Kubie, Lawrence S., 2,133,574
 Kurland, Albert, 310n

 Laennec, R. Th., 15,16
 Lanfeld, E.S., 382
 Langfeldt, G., 137,420
 Laska, Eugene, 540,545,546,547,566,572,574
 Lasky, J.J., 263
 Latent structure model (clustering), 224
 Lazarsfeld, P.F., 207,224,231,262
 League of Nations, Health Section, 93
 Learning approach, 355,357
 Ledley, R.S., 468
 Lehmann, E.L., 236,322,491
 Lehmann, Heinz E., 2,3,22,23,24,72,91,322,329,330,-
 341,346,347,349,491,501,534,553,554,563,571,572,574
 Levine, D., 454
 Levine, Jerome, 575
 Levinson, D., 399n
 Lewis, Nolan D.C., 363,447
 Lindzey, G., 399n
 Linear relations and linearity, 7,165,168,242,346,546,-
 558
 discriminant score, 510
 "Linear Statistical Inference and its Applications"
 (Rao), 561
 Lingo, J.C., 407
 Linnaeus, 5,21,89,135,137,191,201,325,532,547
 Linton, Harriet, 398n
 Lipmann, Ronald, 575
 Lohnes, P.E., 510
 London Psychoanalytic Institute, 62
 "Loosened construction" 161
 Lorr, Maurice, 8,132,225,**261**,262,263,268,269,275,302,-
 308,316,322,323,325,331,336,337,341,342,346,349,367,-
 427,482,495,500,566,571,574
 Lorr Scale, 131
 Lowenfeld, I.E., 370
 Lowenstein, O., 370
 Lowery, Henri A., 300n,574
 Lowrey, L.G., 120
 LSD, 537
 Lubin, Ardie, 243,322,326,327,328,329,344,347,348,-
 403,415,491,500,564,566,571,574
 Lusted, L.B., 468
 Lyster, Samuel B., 1,6,200,213,**220**,240,241,300n,303,-
 305,366,569,571,574
 Lynch, Mary, 35n

 McDonough, M.M., 427
 McGaughan, L.S., 441n
 McLean Hospital, 118,121
 McMillan, John J., 575
 M'Naghten rules, 89
 McNair, D.M., 261n, 262,263,268,269,500
 McNemar, Q., 378,404
 McQuitty, Louis L., 224,243,262,574
 Macfarlane, Jean, 179,182
 MacQueen, James B., 412
 Mahalanobis, P.C., 207,262,283,284
 Mahalanobis distance, 239
 Maintenance of contact perspective, 382
 Malamud-Sands Scale, 131
 Malamud, W., 450
 Manic depressives, 66,81,95,**106**,468
 admission rates for, 105,107
 and schizophrenia, 473
 MAO inhibitors, 474,476
 Maryland, admission rates in, 105,106,**106**
 Maryland Psychiatric Case Register, 59n
 Marzolf, S.S., 132
 Maskin, M., 403n
 Masserman, J.H., 403n
 Mattsson, Eira I., 531

- Mattson, Nils B., 5,315n,349,404,405,**507**,509,518,-532,540,542,545,546,566,572,574
- Maudsley Hospital and Medical Research Council, 107,444
- Maximum fall in systolic blood pressure, 439
- May, J.V., 120,326
- Mayer-Gross, W., 132
- Mayman, M., 137
- Mayo Clinic, 501
- Mayr, E., 118n
- Mazes subtest, 378,382
- Mecholyl, 467
- Medical care data and statistics, **76**,84,85,86,100
- "Medical Orthodoxy and the Future of Psychoanalysis" (Eissler), 133
- Medical Research Council. *See* under Maudsley Hospital
- Medical terminology. *See* "Current Medical Terminology"
- Medicare, 73,112
- Mednick, Sarnoff, 357
- Meehl, P.E., 132
- Meldman, M.J., 131
- Mellaril, 8
See also Thioridazine
- Melzack, R., 357
- Mendelejeff, D.I., 359
- Menninger, Karl, 88,89,117,131,132,137,138,172,352
- Menninger psychotherapy research study, 139
- Menninger Scale, 131
- Mental deficiency, 41,81,104
See also Feeble-mindedness; Mental retardation
- Mental deviance, continuity or discontinuity of, 358
- Mental disease, statistics of, 36
See also Statistics
- Mental disorders, classification of, 36,**99**,101,132
incidence and prevalence of, 100,114
international comparison of, 107
- Mental hospitals, admissions and marital status, 80
data, 99
See also Hospital; Outpatient psychiatry
- Mental illness, classification of, **535**,
as myth, 355
nosology of, **53**
- Mental retardation, 41,91,95,176,178
See also Mental deficiency; under Children
- Mental status, 466
evaluation of, 122
examination, 92,121
measurement of, 197
- Mental Status Schedule, 361,363,364,365,557
- Meprobamate, 477
- Methacholine, 535,537
- "Method of Multi-working Hypotheses, The" (Chamberlin), 135
- Meyer, Adolf, 92,98,121,344,419,420,421,422,424,503
- Meyer, Fred, 286
- Meyerowitz, Sanford, 35n
- Michael Reese Hospital, 250
- Michaux, Mary, 306n,310n
- Michigan schizophrenia project, 5,405,408
- Michigan, University of, 5
- Miles, Harold, 35n
- Minnesota-Hartford Personality Assay, 484,485,486,-487
- Minnesota Multiphasic Personality Inventory. *See* MMPI
- Minnesota, University of, 435
- Mixture problem (in statistics), 197,199,200,559,-564,565
- MMPI, 404,406,407,410,435,**436**
- Mongolism, 83,88,91
- Monro Scale, 131
- Monroe County (N.Y.) Psychiatric Case Register, 4,40,41,42,42,44,46,48,49,50,59n,106
- Monroe, Russell, 535,575
- Moore, T.V., 131
- Moran, L.J., 441n
- Morrow, J. Tarlton, 180
- Morgagni, G. B., 14
- Motivation studies, 375
- Motor disturbances, 263,264
- Mowrer, O.H., 136n
- MRC. *See* Maudsley Hospital and Medical Research Council
- MSS. *See* Mental Status Schedule
- Multidimensional scaling, 218
- Multiple correlations, 496
- Multiple cutting scores, 270
- Multiple diagnoses, 75,98,148
- Multiple discriminant analysis, 283,288,518
- Multiple-regression approach, 258,479,482,504,564
- Munchausen syndrome, 93
- Murphy, Lois B., **173**,188,540,543,544,571,574
- Murray, H.A., 136,137,391,404
- "Naming is knowing", 151,189
- National Association for Mental Health, 37n
- National Clearinghouse for Mental Health Information, 59
- National Institute of General Medical Sciences, 410n
- National Institutes of Health. *See* NIH
- National Institute of Mental Health-Psychopharmacology Service Center. *See* NIMH-PSC
- National Library of Medicine, 60n
- Needs (human), 391,395
- Negroes, 417,544
- Nervous Child (publication), 176
- Neuhaus, J., 511
- Neulinger, John, 173,183,**390**,397,398,403,404,413,-414,415,416,572,574
- Neurophysiological approach to classification, 355,-358
- Neuropsychology and neuropsychopharmacology, 179,501

- Neuroses, 41,48,62,66,145
 and psychoanalysis, 68,70
See also Institutional neurosis; Organ neurosis;
 Pseudoschizophrenic neuroses; Psychoneurosis;
 Transference neuroses
- New York State Hospital system, 121
- New York State Psychiatric Institute, 107
- New York University group and research, 392
- Newman-Keuls procedure, 314
- NIH, 127n,328,346,501
- NIMH, 37n,59,105,263,363
 —APA Conference Planning Committee, 569
 drug response study, 480,483
See also NIMH-PSC Collaborative Study
 phenothiazine study, 315,316,363,416
 —Psychopharmacology Research Branch, 144
 —PSC Collaborative Study of Drugs and
 Schizophrenia, 104,286,290,291,292,295,307,-
 308n,313,337,340,342,495,497,498,502
- Nomenclature and nomenclatures, 95
 and statistical classification of diseases, 100
 vs. classification, 148
See also under name
- Nonfragmentation, 165,167
- Nonlinearity, 7,9,165,168,328,346,546,553,558
- Norepinephrine, 513,529
- Normalized vector product model, 284
- Nosology, 53,95
- Nosology Conference (1965), 60,61
- "Nosology of Schizophrenia, The: A Cooperative
 Study" (R.W. Gerard), 508
- Nosology and Technical Vocabulary Review Com-
 mittee, 60,61
- Numerical taxonomy problem, 508
- Nunnally, Jum C., 209,224,225,249,262,322,323,326,-
 335,336,341,345,350,351,410,571,574
- Nystagmus, 536
- Object Assembly and Object Classification Tests,
 382,442-443
- O'Connor, J.P., 302,379
- Observation and observer, criteria for, 140
 reliability and variability of, 102,109
- Obsessive-compulsives, 64,66,67
 and depression, 470,471
- "On the Nature of Control in the Psychological
 Experiment" (seminar), 116
- "One person, one disease," 4-5,39,84,94,98,99,242
- Optimum decision rule, 230
- Order statistical method, 508
- Organ neurosis, and psychoanalytic treatment, 62-63
- Osgood, C.E., 205,284
- "Outlines of Psychiatric Examination" (Lewis), 363
- Outpatient psychiatry, 35,37,39,47,99
 admissions, 46
 classification in, 35,50-51
 diagnostic terms and, 40,41,42
- Outside contamination, 233
- Overall, John E., 5,8,132,173,183,205,206,207,209,277,-
 278,280,281,282,283,287,296,322,323,324,325,326,328,-
 331,338,342,346,350,351,482,566,571,574
- Oxypertine, 295,297,324,389
- Panic and panic states, 91,469,470,476,477
- Paracelsus, 21
- Paranoid and paranoids, 43,52,66,67,90,91,127,263,-
 264,297,339,468,497,498
 interpersonal disturbances, 298,299
 and triperidol, 324
- PAS. *See* Personality Assessment System
- Pasamanick, Benjamin, 3,24,36,74,81,149,150,189,-
 356,540,543,544,572,574
- Passive-aggressive character disorders, 473
 and personality, 43,64,66
- Pathology, 12
 autonomous and drug-responsive, 483
See also "Standardized Nomenclature on Path-
 ology"
- Patient and patients, biometric assessment of, 353
 characteristics, 3,8,10,99,274,328,420,507,565
 classification of, 99
 as an entity, 161
 types, 6,499
 and drug treatment response, 314,315
- "Patients in Mental Institutions" (NIMH), 104n
- Pavlov, I.P., 331,478
- Payne, R.W. and Payne Object Classification Test,
 441n,442
- Peace Corps group, and research, 392
 types, 400
- Pearson, Karl, 200,227,262,394
- Pennington, Veronica, 286
- Pentothal, 535,536
- Pepernik, M.C., 425,427
- Pepitone, A.D., 398n
- Perceptual distortions, 263,264
- Perceptual isolation, 535
- Perceptual tests, 359
- Perceptual threshold, 382
- Pereira, Jonathan, 17
- Perphenazine, 289,295,297
- Personality, 345
 and psychopathology, 501,505,559
- Personality assessment, 559
- Personality Assessment System, 377,379n,380,382,-
 387n,404
- "Personality Change" (Byrne and Worchel), 344
- Personality disorders, 24,41,43,48,49,50,52,64,66,67,-
 68,104,134
 admission rates for, 104
 and hospitalization, 45,49
- Personality patterns and structure, 102,399
- Personality rating scales, 122
- Personality tests, 8,130

- Personality theory, 5,157,159,161,165,166,167,402
 classification in, **155**
 Personality trait disturbances, 43,134
 Perversions, 66,68,70
 Pharmacodynamics, Galenic, 13
 Phenothiazine, 143,289,467,471,472,473,475,477,478,-
 482,504,537n
 and panic states, 8
 study (NIMH), 268,315
 Phenylalanine, 357
 Phenylketonuria, 24,89,192,243,357,501,535,556
 Phillips, L., 131,142,430,454,455,456
 See also Phillips Scale
 Phillips Scale of Premorbid Adjustment in Schizo-
 phrenia, 426,429,430,431,432,433,435,441,444,447,-
 450,453,453,**464**,495,499,502,503,505
 Phobia and phobic, 64,66,90
 —*anxious* states, 469
 "Phobias and Their Vicissitudes" (psychoanalytic
 study), 67
 Physiology, 12,17,355,535,543
 Piaget, Jean, 449
 Picture Arrangement and Picture Completion Tests,
 378,382
 Pieper, Samuel J.L. Jr., 57n
 Pilgrim State Hospital, 75
 Pinel, Philippe, 13,14,15,16,18,29,354
 Pisoni, Stephanie, 367
 Plag, John, 326
 Plant, J.S., 121,131
 Plato, 136
 Pollack, M., 474
 Pollin, W., 131
 Pollock, Horatio, 36
 Populations, classification of, 77,83
 Porterfield, J.L., 281
 Portnoy, Stephanie, 367,368
 Prediction of Change Scale, 435
 Premorbid adjustment scale, *See* Phillips Scale
 Prescott, James, 575
 Prevalence rate, 76,77
 Prevention, 39,54
 Prien, Robert, 575
 Principal components analysis, 511,561
 Probability theory, 231,282
 "Probability Theory and Psychoanalytic Research"
 (Chassan), 62
 Profile and profiles, 195,204,278,285,293,294
 analysis techniques, 277,287
 clusters, 255
 depressive, 297
 elevation, 208
 paranoid, 297
 schizophrenic, 297
 Prognosis, 419,424,428,505
 See also Prognostic Scale for Schizophrenia
 Prognostic Scale for Schizophrenia, 435
 Projective techniques, 407,418
 Proverb test. *See* Benjamin Proverb Test
 Pruyser, P., 137
 PSC, 502
 Pseudodichotomous variables, 326
 Pseudoschizophrenic neuroses, 476
 Psychiatric and psychiatry, 1,39,75,87,88,99,112,113,-
 149,158,161,328,440,450,466,467,549,551,553,554,556,-
 563,567
 behavioral and medical aspects of,133,135,153
 case registers, 40,44-45,105
 See also under name
 classification in, 10,76
 examination, 363
 hospital. *See* Hospital psychiatry; Outpatient
 psychiatry
 nosology, **190**
 social, 107
 subclinical cases in, 81
 Psychiatric Status Schedule, 365
 Psychoanalysis and psychoanalytic, 4,**62**,127,132,133,-
 174,555
 diagnosis in, 68,71
 vs. psychotherapy, 70
 statistics, 62
 study of result evaluation, 68
 "Psychoanalysis and Human Values" (Masserman),
 403n
 Psychological and psychology, 116,117,134,157,161,-
 166,167,195,328,418,549,551,556,567
 abnormal. *See* Abnormal psychology
 —*mindedness*, 382
 social, 170
 tests, 406,467,558,565
 prognostic value of, 415
 See also under name
 See also Clinical psychology, Gestalt psychology;
 Neuropsychology
 "Psychological Types" (Jung), 320,321
 Psychometrics, 164,197
 Psychoneurosis and psychoneurotic, depressive re-
 actions, admission rates of, 105,106,106
 disorders, 41,43,470
 admission rates of, 104,105,106
 and psychoanalysis, 62
 See also Neuroses
 Psychopathology, 8,24
 classification of, 10,76,77,**116**
 and personality, 501,505,559
 Psychopharmacology, classification in, 10
 "Psychopharmacology and Personality" (Zubin and
 Katz), 344
 Psychopharmacology Service Center. *See* PSC
 Psychophobia, 434
 Psychophysiological disturbances and disorders, 41,-
 43,64-65,67
 Psychosexual development, 505

- Psychosis and psychotic, 66
 functional. *See* Functional psychosis
 and hospitalization, 45,275
 and psychoanalysis, 63
 psychogenic, 43,104,192
 schizophreniform, 420
 types, 271,272,273,275,276
 acute, 273
 chronic, 267,272,275
 long-term, 267
 men, 265
 women, 266,272
See also Affective psychosis; Arteriosclerotic
 psychoses
- Psychosomatic disorders, 65,134
 and psychoanalysis, 63
- Psychotherapy, 70,159,164
- Pupillography, 370,376,535
- Pythagorean distance, 206,279,283,285
- Q-factor analysis, 8,9,251,252,304,305,327,392,402,414,-
 509,511,512,513,514,515,518
- Quantitative approaches, **277**
- Rabinovitch, M.S., 131
- Raines, George, 36
- Rank, Otto, 403n
- Rao, C. Radhakrishna, 1,7,9,199,213,214,220,226,**229**,
 235,240,241,242,243,245,246,288,510,549,551,559,561,-
 571,574
- Rapaport, D., 383,404
- Raskin, Allen, 575
- Ratings, by nursing personnel, 311,483,484,485,486,-
 487,489
 psychiatric, 308,309,309,310,467
 by relatives, 5,8,300,304,306,307,312,340
 validity of, 312,335
- Raw score analysis, 255
- Recording vs. classification, 74,75
- Rees, L., 202
- Regression and regressive equations, 175,502
See also Multiple regression
- Reid, D.D., 102
- Reiss-Davis Clinic for Child Guidance, 56n
- Relatives' ratings. *See* Ratings by relatives
- Remissions, 174,468,469,480,481
- Repetitive sensory stimuli, 535
- Research, **116**
 classification in, **143**
- Research reports, 247
- Retardation, and apathy, 263,264
 and imipramine, 324
- Rice, Charles E., 507,518,575
- Rimland, Bernard, 176,575
- Rochester, University of, 40
- Rockland, L.H. and Rockland-Pollin Scale, 131
- Rodnick, Eliot H., 432,434,440,448,452,531
- Rogers, D.J., 206
- "Role of Statistics and Psychoanalysis" (Chassan),
 62
- Rorschach Test, 408,409,415,416,417,438,439,449,454,-
 537n
- Rosenberg, M.J., 397-398
- Rosenthal, D., 443,444,445
- Rosenzweig, Saul, 127,130,136n, 140n
- Rotation criteria, 511
- Roth, M., 108,535
- Rumke, H.C., 301
- Sainz, Anthony, 500
- Saliency, 257,350
- Salzinger, K., 367,368
- Sampling, 504,531
- Sandifer, Myron Jr., 75,575
- Sands, S.L., 131
- Sargant, W., 475
- Sargent, Helen, 139
- Santayana, George, 133
- Saunders, David R., 8,226,262,277,326,**377**,378,379,-
 382,388,403,404,508,572,574
- Sawrey, W.L., 225,226,262
- Scaling, 362
- Scatter, 208,208
- Schaefer, S.M.J., 435,436
- Schaefer, Theodore, 356
- Schanberger, W., 427
- Scheff, Thomas, 172
- Schizoaffective reactions, 43,468
- Schizoid personality, 66,67
- Schizophrenia and schizophrenic, 5,49-50,81,161,-
300,475,**419**,494,**507**
 acute, 8,268,273,300,367,370
 admission rates for, 105
 childhood. *See* under Children
 classification of, 121
 See also Schizophrenic types,
 classification of,
 chronic, 43,367,370
 communication in, 368
 diagnosis of, 46,47,48,49
 disorganization syndrome, 495,496,497,498
 —drug study, 307,313,317
 family studies, 408
 genetic, 6,443,445
 and hospitalization, 38,49,49,359
 latent, 421
 Michigan study of, 5,405,408
 MMPI profiles, 436
 and neurosis, 184
 and nonschizophrenics, 9,369,511,533
 paranoid, 8,9,43,468
 of personality disorder, 50
 phenomenological typology of, 300
 and phenothiazine, 143,268,315
 process-reactive concept of, 9,419,426,438,446,-
 503,504,505

Schizophrenia and schizophrenic—Continued

- prognosis in, 427
- pseudoneurotic, 67,470
- and psychoanalysis, 68
- psychological concept of, 68
- reactive, 5,6,9,491,492,494
- size estimates of, 448,449,449
- types, 510,515,529,530-531,533
 - classification of, 520
 - discriminant analysis of, 519,523,525
 - typology of, 300,**507**
 - undifferentiated, 43
- See also* Group of schizophrenias
- Schmiedeberg, Oswald, 17
- Schofield, W., 435
- Schucman, Helen, 226,262
- Schofield, W., 435
- Science, 54,156,157,563
 - classification in, 117,160,191
 - psychosocial, 133
- Science (publication), 135
- Sedation threshold studies, 535,536
- Seidel, Claudine, 430n, 432
- Self-Description Questionnaire and self descriptions, **390**,391,392,396n,397,400,402,418
- Self-report inventories, 406,418
- Sequential decision rules, 233
- Sexual disorders, 41,64,67
 - and psychoanalysis, 62
- SGOT test, 552
- Shagass, C., 371,535
- Shakow, David, 1,**116**,127,133n,147,151,152,153,154,-172,174,352,403,405,416,417,450,451,452,566,570,574
- Shanan, Joel, 399n,400n
- Shatzky, Jacob, 90
- Shawver, J.R., 278
- Shelton, Jack, 286
- Shepard, R.N., 214n
- Sherif, M., 398n
- Silverman, J., 451
- Similarities and similarity, 294,378,382
 - index, 205,211,517
 - between people, 6,200,**201**,221,511
 - between profiles, 195
 - structures, multidimensional representation of, **212**
- Simkin, J.S., 378
- Simpson, George, 36,118n
- Singer, Margaret T., 408,416
- Situational tests, 418
 - See also* Transient situational disturbances
- Slater, E., 443
- Slater, Patrick, 239
- Sleep-threshold studies, 535,536
- Slotnick, Lee K., 439
- Smith, C.A.B., 235
- Smith, D.E.P., 382

- Smith, Edwin. *See* Edwin Smith
- SNDO. *See* "Standard Nomenclature of Diseases and Operations"
- Sneath, P.A., 239
- Social Adaptation Inventory and Anamnesis, 363
- Social Adaptation Schedule, 365n
- Social Adjustment Scale, 113
- Social breakdown syndrome, 82,93-94,503-504
- Social Competence Index, 456
- Social sciences, 134
 - See also* under name
- Socialization Scale, 404
- "Sociobiological Types and Methods for Their Isolation" (Zubin), 557
- Sociocultural approach, 47,49,179,355,356,358,417
- Socioeconomic status, 47,49
- Sociology, 5,24,78,179,551,567
 - classification in, **168**
- Sociopathic personality, 43
- Sokhal, R.R., 239
- Solomon, L., 433,433,435,439
- Somatopsychic disorders, 134
 - See also* Psychosomatic disorders
- "Some Uses and Abuses of Statistical Methods in Psychiatry" (Jellinek), 129
- Sonn, Thomas, 575
- Southard, E.E., 119
- Spence, Janet T., 441,442
- Spitzer, Robert L., 361,363,375,415,575
- Spitzer Scale, 131
- Split diagnoses, 511
- Spring Grove State Hospital, 303,306,310n
- Stainbrook, E., 132
- Standard diagnostic nomenclature, 280
- "Standard Nomenclature" (APA), 37,114,148,286,-293
- "Standard Nomenclature of Diseases and Operations", 100,114,150
- "Standardized Nomenclature on Pathology", 148
- Standards in psychiatry and psychology, 140
- Starkweather, J.A., 405
- Statistics and statistical, 6,129,166,**197**,227,332,342,-354,490,549,553,554,559,560,563,565,567
 - of classification, 195
 - definition, 353
 - international, 103
 - methods, 129
 - sorting, 508,511,516
- Stearns, A.W., 119
- Steffy, R.A., 428
- Stein, Morris I., 9,173,183,**390**,400n,403,404,406,413,-414,415,417,572,574
- Stengel, E., 36
- Stephenson, W., 262,304,509n
- Stereotypes, clinical and psychiatric, 225,339
 - cultural, 188

- Stimulus variables and activities, 360,361
- Strecker, Edward M., 123
- Strong, E.K. and Strong Vocational Interest Blank, 381,404
- Structured Clinical Interview, 363,364,365,557
- Successive factorial screening, 224,518,532
- Suci, G.J., 205,284
- Suicide attempts, 41,48
- Sullivan, Harry Stack, 403n,419,422,424
- Sutton, S., 359,360,369,370
- Sydenham, Thomas, 13,555
- Symptom and symptoms, clusters, 330
complexes, 90
patterns, 102
primary and secondary, 559
and psychoanalytic study, 66
See also Syndromes; Target symptoms
- Syndrome and syndromes, 91,93,263,552
analysis, 508
classification of, 90,91
clusters, 90,286
delineation methods, 552
new, 81
psychotic, 337
scores, 271,272
See also under name
- Synthetic approach to clustering, 221,222
- Systematics and new systematics, 118
- Szasz, Thomas S., 172
- Tanimoto, T.T., 206,508,509,517
- Target symptoms, 477,478
- TAT, 408,416,417,446
- Taxonomy, 117,166,353
- Temkin, Oswei, 2,11,20,21,22,23,24,25,75n,570,574
- Test and tests, culture-related, 8
intelligence. *See* I.Q.
social-class related, 8
See also under name
- "Textbook of Psychiatry" (Bleuler), 440
- "Textbook of Psychiatry" (Henderson and Gillespie), 450
- Thematic Apperception Test. *See* TAT
- Theory, 155,156
function and structure of, 162,163
- Therapeutic response, evaluation of, 3
prediction of, 3
See also Treatment response
- Thinking disturbance syndrome, 298,299
- Thioridazine, 8,288,289,295,297,314,317,317,329,495
and anxiety, 324
and schizophrenia, 340
- Thompson, W.R., 357,361
- Thorason, G., 509n
- Thorndike, R.L., 227,262
- "Thugs of India", 149
- Thurstone, L.L., 378
- Times (in epidemiology), 77
- Tofranil, 8
See also Imipramine
- Topeka Psychoanalytic Institute, 62
- Torgerson, Warren S., 1,3,212,214n,240,241,242,298,-482,545,552,564,571,574
- Training and training programs, 133
vs. education, 54
- "Training the Psychiatrist to Meet Changing Needs" (GAP), 259
- Trait dimensions and traits, 250,251
- Transference neuroses, 62
- Transient situational personality disturbances, 43,-51,95,148,149
- "Treatise on Insanity" (Pinel), 29
- "Treatise on the Principles and Practice of Medicine" (Flint), 16
- Treatment evaluation and classification, 32
- Treatment goals, 82
- Treatment response, 244,324,325
prediction of, 8,143,564
relation to pretreatment symptom, 289
typology, 325
- Tri-County Psychiatric Case Register, 59n
- Trifluoperidol, 289,297
- Triperidol and paranoids, 324
- Tryon, R.C., 407
- Tucker, L.R., 509
- Tuke, Daniel H., 2,90,116
- Tulane University, 56n
- Tuma, Hussein, 575
- Tutko, T.A., 441,442
- Twin studies, 443,445
- Type and types, 243,264
action oriented, 395,404
—definers, 393
definition, 243
—by-drug-interaction, 317,323
intellectually oriented, 395,404
Peace Corps, 392,400,414
reliability and validity, 269
resourceful, 396,404
socially oriented, 395,404
sociobiological, 557
unconventional, 396,404
See also "Psychological Types"
- Typologies, 247,256
- Ulmann, L.P., 431n
- U. K., committee on vital and health statistics, 102
Subcommittee on Classification of Psychiatric Disorders, 112
- U. S. Army classification, 93
- U. S. National Committee on Vital and Health Statistics, 101-102
- "Use and Interpretation of Principal Components Analysis, The" (Rao), 561

- VA Cooperative Studies in Psychiatry, 8,270,495,496
 VA Nomenclature, 37
 VA psychiatric evaluation project, 431
 VA schizophrenia-drug studies, 286,347,480
 Verdun Protestant Hospital, 332
 Veterans Administration. *See* VA
 Virchow, R., 16
 Visotsky, Harold M., 73,575
 Visual evoked cortical response, 535
 Visual flicker fusion, 535
 "Vital Balance, The" (K. Menninger), 88,137,172,-
 352
 Vocabulary Subtest, 382

 WAIS, 378,387,446
 Wallner, J.M., 425,426,427,438,439
 Ward behavior, 8,121,138
 Ward Behavior Inventory, 363,365
 Ward Behavior Scale, 311,311,312,504
 Ward, J.H. Jr., 206
 Waskow, Irene, 575
 Waste-basket diagnoses. *See* Diagnoses, waste-basket
 WBI. *See* Ward Behavior Inventory
 Webb, W.W., 431
 Wechsler Adult Intelligence Scale. *See* WAIS
 Wechsler-Bellevue Intelligence Tests, 8,387n,404
 Wechsler, David, 377,378,379,381,383,387,388,404
 Weinstock, Harry I., 1,4,62,72,73,74,164,166,173,184,-
 555,570,574
 Welch, B.L., 235
 Welsh, G.S., 435
 Werble, Beatrice, 252n,260,575
 Wernicke's encephalopathy, 87,91
 Westphal-Edinger nucleus, 370
 Wherry, R.J., Jr., 326
 Whewell, William, 89
 White, W.A., 123
 WHO, 35,100,102,112,148

 Wiener, H., 425
 Wilson, E.B., 119,120
 Wilson, Paul T., 53,575
 Winder, C.L., 425,426,427,438,439
 Windle, Charles, 415
 Winer, B.J., 314,314n,317n
 WISC, 378
 Withdrawal (psychic), 175,185
 and drug response, 317
 —retardation syndrome, 298,299
 Wittenborn, J.R. and Wittenborn Scale, 131,302
 Wittman, Phyllis, 426,427,428,429,437,437,438,453,-
 502
 Wittson, C.L., 132
 Woodbury, Max A., 572,574
 Woods, Joan B., 57n
 Worcester State Hospital, 1,116,118,119,120,121,127n,-
 130,131,136,139,140,151,172,450,451,452,453,459
 Worchel, P., 344
 World Health Assembly, 100
 World Health Organization. *See* WHO
 "World Health Organization Classification" 90,93
 World War II, 36,37
 Wrigley, C., 511
 Wynne, L.C., 408,445

 Ypsilanti State Hospital, 5

 Zalusky, M., 432
 Zigler, E., 454,455,456
 Zilboorg, G., 419
 Zimet, C.N., 449
 Zlotowski, M., 433,433,435,439
 Zubin, Joseph, 3,8,9,24,75,227,241,262,305,322,344,-
 353,356,359,360,363,367,369,373,403,404,405,412,413,-
 415,416,417,491,500,505,508,549,556,566,567,572,574
 Zubin Scale, 131

DATE DUE

~~NIH LIB APR 9 1973~~

~~JUN 26 1981~~

~~APR 25 1981~~

GAYLORD

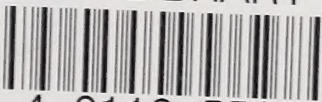
PRINTED IN U.S.A.



<http://nihlibrary.nih.gov>

10 Center Drive
Bethesda, MD 20892-1150
301-496-1080

NIH LIBRARY



4 0118 5525

NIH LIBRARY



3 1496 00268 7435

~~SEP 1 1987~~

Public Health Service Publication No. 1

U.S. DEPARTMENT OF HEALTH, EDUCATION, AND

Public Health Service